# JOURNAL OF APPLIED PHYSICS

JOURNAL
OF
# APPLIED PHYSICS

NOVEMBER, 1942 Vol. 13, No. 11

## *In This Issue*

*Page*

**Physics in Industry** E. U. CONDON 663

**Color Television**
P. C. GOLDMARK, J. N. DYER, E. R. PIORE, AND J. M. HOLLYWOOD 666

**Physical Methods of Dosage Determination in Radiation Therapy** EDITH H. QUIMBY 678

**Frictional Phenomena. XI** ANDREW GEMANT 688

*Contributed Original Research*

**Rapid Infra-Red Analytical Method for Hydrocarbon Mixtures and a Routine Spectrophotometer for Plant Control** R. ROBERT BRATTAIN AND O. BEECK 699

**Steady State Currents of Electrical Networks** D. L. WAIDELICH 706

**Formulas for the Inductance of Rectangular Tubular Conductors** THOMAS JAMES HIGGINS 712

**Fatigue of Fabrics**
W. F. BUSSE, E. T. LESSIG, D. L. LOUGHBOROUGH, AND L. LARRICK 715

**Specific Surface Measurements on Compact Bundles of Parallel Fibers** R. R. SULLIVAN 725

*In Every Issue*

**New Books** 687

**Here and There** 697

**Calendar of Meetings** 698

**Index to Advertisers** xii

**BOARD OF EDITORS**

ELMER HUTCHISSON, *Editor*

*Associate Editors*

FRANCIS BITTER
W. F. BUSSE
S. H. CALDWELL
ROBLEY D. EVANS
MORRIS MUSKAT
FREDERICK SEITZ
B. E. WARREN
J. B. WHITEHEAD

**Publication of the American Institute of Physics**

**ADMINISTRATIVE STAFF**

HENRY A. BARTON, *Director*

MADELINE M. MITCHELL, *Publications Manager*

JOHN T. TATE, *Adviser on Publications*

**GOVERNING BOARD**

PAUL E. KLOPSTEG, *Chairman*
GEORGE B. PEGRAM, *Secretary*
R. B. DOW
L. A. DUBRIDGE
FLOYD A. FIRESTONE
HARVEY FLETCHER
R. C. GIBBS
K. S. GIBSON
ARTHUR C. HARDY
GEORGE R. HARRISON
HERMANN MARK
A. G. WORTHING
JOHN T. TATE
N. W. TAYLOR
WALLACE WATERFALL

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. It aims to be of service not only in physical laboratories, but also in the laboratories of industrial, chemical, geological, geophysical, meteorological, radio, and similar concerns. The JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists and reviews, as well as technical papers of applied physics.

**Manuscripts** should be submitted to Elmer Hutchisson, Editor, American Institute of Physics, 175 Fifth Avenue, New York, New York.

**Subscriptions** and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 175 Fifth Avenue, New York, N. Y.

SUBSCRIPTION PRICE

| | U. S. AND CANADA | ELSEWHERE |
|---|---|---|
| To members of scientific societies | $5.00 | $5.70 |
| To all others | 7.00 | 7.70 |

**Single copies** are $.70 each.

**Changes of address** and proofs should be addressed to the Publications Manager.

**Advertising rates** supplied on request. Orders, advertising copy and cuts should be sent to the American Institute of Physics.

Entered as second class matter at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1897. Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1942, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 13 November, 1942 Number 11

# Physics in Industry*

BY E. U. CONDON

*Associate Director, Research Laboratories, Westinghouse Electric and Manufacturing Company, East Pittsburgh, Pennsylvania*

WE are assembled here tonight to dedicate to public service a splendid new addition to the plant of Purdue University. We have the honor of opening what I believe to be the first university laboratory of *applied* physics in America—a generous gift to Purdue in memory of Charles Benedict Stuart.

We do this at a time when the nation's universities and the nation's industries are working together as never before to do their part in the battle for freedom. It is fitting at such a time that we look over that part of civilized life for whose cultivation we as physicists are responsible, that we try to see how our science has been developing and in what direction we may cause it to grow in the future. We want to see clearly the job ahead so we can set about doing it earnestly and cheerfully, for we believe that, with effort, we can determine the future in peace as in war.

Although we can trace contributions to physics in America as far back as, say, Benjamin Franklin and his kite, the reasonably widespread organized research in physics that is now such an important feature of American academic and industrial life is a product of the last twenty-five years.

Our professional group, the American Physical Society, is not yet fifty years old. For the first twenty years or so of its existence, very little was reported by its members which has affected the later development of our science in a basic way. The period was one of thoroughly sound beginnings marked by a few really outstanding contributions: Rowland discovered the magnetic effect of moving electric charge and he perfected the diffraction grating for spectroscopy; Hall discovered the Hall effect which is today a valuable tool in studying electrical conduction in metals and crystals; Michelson and Morley discovered the phenomena which laid the groundwork for the theory of relativity. Somewhat later Millikan measured the charge on the electron and the Planck constant.

It was a period in which very few of our physicists had the opportunity to contribute much to the advancement of science. Most of them were concerned with teaching, and the great industrial laboratories as we know them today simply did not exist.

A great change set in just after the first World War, although it is hard to see that that had much to do with it directly. Indirectly it did, for war acts as a stimulus to cooperative effort. The National Research Council was organized and soon after the war the Rockefeller Foundation established the National Research Council Fellowships in physics, chemistry, and mathematics. By this means some fifty young Ph.D.'s were each year enabled to continue their de-

* Address delivered at the opening of the Charles Benedict Stuart Laboratory of Applied Physics, Purdue University, June 19, 1942.

velopment as research men to such a degree of self-reliance that most of them could shoulder other burdens without losing their ability to carry on in research. I believe that these Fellowships alone, supported at a cost that was absolutely trifling, have done more for the development of American research science than any other single thing. The Fellowships reacted not only on the Fellows but on the institutions and the faculties at the universities where the Fellows worked.

In the 1920's research in fundamental physics expanded by leaps and bounds. The momentum carried by x-ray quanta was demonstrated by Arthur Compton; Davisson and Germer discovered electron diffraction; investigation of cosmic rays was inaugurated; atomic and molecular spectra were analyzed; and a group of young American theoretical physicists participated vigorously in the development of quantum mechanics, and so on.

It was also a period of rapid expansion of industrial physics in the electronic industries. From the birth of radio broadcasting in Pittsburgh in 1921 there was steady progress in the development of useful applications for phenomena associated with electric discharge in gases. The applied science of electronics was born to serve as a strong bond between physics and electrical engineering.

However, it was rather characteristic of that period that industry and physics were not closely associated. Academic physicists were a little inclined to be mistrustful of industry. In some vague way it was felt by many that the interests of engineering and of physics were best served by keeping them far apart after the engineering student had finished his sophomore course in general physics! Many a bright young man, turned physics instructor, felt that he was casting pearls before swine when he had to expound the beauties of his subject to the sophomore engineers. Oddly enough, the engineering students could detect this air of condescension in their physics instructors and did not react to it with humility. I am often surprised at the vehemence with which leading engineers whom I meet will denounce the physicists who taught them in college. It was a most unhealthy situation, happily now almost completely changed.

Then came the 1930's, largely a shameful decade in which people lived in insecurity, with much unemployment, little hope of professional advancement, and in which there was a totally inadequate development of our natural resources in scientific talent and inventiveness. Like ignorant fools we watched while our present enemies bent every effort toward the mobilization of their universities and their industries to prepare the means that would crush us as a free people. In 1934 when Hitler started to rearm, many of our physicists were unemployed, and our young people could see little opportunity for a career in science. By 1936 the armament program had progressed so far in Germany as to require speed-up in the universities, such as we now have, in an attempt to supply the vast numbers of technical men which modern warfare requires. But we slept on, seeing in this not a marvellously coordinated effort to conquer us, but simply viewing it as a distant anti-intellectual attack on their universities. As late as 1940, even in 1941, there were those who argued against taking steps to resist these murderers of Lidice—because to do so would call for disturbing our way of life, and upsetting the calm, detached pursuit of knowledge in our universities!

On top of all this blindness to the international scene, we had to hear attacks on the scientific method as anti-social. Economic and social ills of society culminating in the depression were somehow blamed on science—as if all our troubles arose from too much of that calm analysis of carefully checked observations that is characteristic of scientific method! Somehow it just couldn't be admitted that there were some unsolved problems in applied social science. Somehow the fault lay with the physicists, for having taught us too well to understand the forms of matter and energy and how to use this knowledge to improve man's physical conditions of life. Poppycock!

Silly as such talk sounds now, it had quite a following in the 1930's. Support of scientific research was unduly curtailed except toward the very end of the decade. This situation might have been much worse if it were not for the fact that those physicists who had jobs, in spite of depression conditions, were having such a gloriously good time.

Heavy hydrogen was discovered, the neutron, the positron, artificial disintegration of the elements, artificial radioactivity; isotopic masses were measured with precision; new means for producing high voltage were developed; and finally there came the application of the new knowledge, while not yet five years old, to give medicine and biochemistry new tools of investigation that were soon to make vast changes in their approach to basic problems. With all this going on, is it any wonder that physicists did not find time even to answer the charge that they were responsible for the depression!

To be sure, some of them would have felt better to have had decent jobs, others would have liked, or at least their wives wished for, a little raise. And lots of fine young minds who would be most welcome in the profession today looked elsewhere for work, seeing how limited were the opportunities in science.

But the outlook was not entirely as dark as I have painted it. Toward the end of the 1930's there developed in America a new realization of the fruitfulness of bringing the newest results of physics into industry. Industrial research, already strongly established in the chemical field, began to expand in the direction of applied physics. The more progressive engineering schools established courses in applied physics. More physicists began to find their places in industry. Physicists began to see that their science, far from being restricted or polluted by an association with technical processes, could derive a new stimulus and a new significance that it could never possess so long as it was the private intellectual pursuit of a cloistered few.

This trend found expression in the development of a flourishing research periodical, the *Journal of Applied Physics*. For a time it looked as if there might be a separatist movement whereby the industrial physicists would have professional societies of their own and would have little to do with the academic or "pure" physicists. Fortunately all such tendencies have been eliminated by the policy of the American Institute of Physics in giving due recognition to the importance of industrial physics.

This war has completed the job of bringing all American physicists to a clear understanding of how much their science can contribute to the welfare of the people. At the moment their energies are devoted to applying their knowledge to winning the victory with a minimum of human suffering and material waste. Many university men are learning for the first time that there is a thrill and a deep satisfaction not only in discovering a new principle in science, but also in putting principles into application in new ways. When this war is over they will not forget this experience. Returning to the pure science laboratories in the universities they will have a sympathetic understanding of the problems of industry which will broaden and enrich the relationships of physics and engineering. I think they will even be more friendly toward and more tolerant of the sophomore engineers.

At the same time we must resolve not to neglect the cultivation of the basic science which we hope some day to apply. More and more, industry in America is recognizing the debt it owes to fundamental science—a debt it can hope to repay by fostering more basic research in its own research laboratories and by working in close cooperation with the universities.

I feel sure that those who are entrusted with furthering scientific research in Purdue see this problem of applied physics in all its broad implications. They recognize, as we do in industry, that all physics is applied physics—so-called pure physics being simply that part whose application is to satisfy the curiosity of the physicists. Those of us here tonight who have had the opportunity of inspecting these laboratories today are happy to report that there is ample provision here in Purdue for the cultivation of this kind of applied physics, too. And a fine thing it is, for that is where all the rest of applied physics comes from. Our best wishes to the faculty and students who are to work in this splendid laboratory.

# Color Television*

BY P. C. GOLDMARK, J. N. DYER, E. R. PIORE, AND J. M. HOLLYWOOD
*Columbia Broadcasting System, New York, New York*

## INTRODUCTION

MUCH of the significance of color in television is striking, even to the casual observer. Aside from the most obvious effect, namely, that color introduces a sense of reality and a lifelike quality into the picture, comparison of a color-television picture with the corresponding black-and-white image makes it apparent that not only are small objects more perceptible but outlines in general seem to be more clearly defined. As has been experienced with other media, color in television also seems to introduce a certain perception of depth. This is partly due to the increased ability of color to reproduce the contrasts and shadows as well as high lights and reflections in different hues, while the degree of color saturation, which is a function of distance, especially outdoors, strongly enhances the three-dimensional quality.

Effects such as those mentioned here became apparent immediately after initial experimentation with color television and were encouraging enough to warrant an extensive investigation of that field with the objective of producing a practical color-television system.

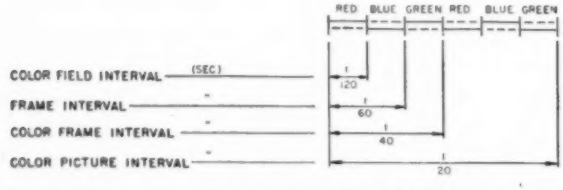



FIG. 1. Diagrammatic representation of CBS System 3.

At the outset it was realized that the transmission and reception of live objects as well as motion-picture film in natural color entailed the use of a trichromatic system.

Color television was broadcast for the first time on August 28, 1940, over station W2XAB of the Columbia Broadcasting System in New York City. The system used was a three-color, high-definition television system employing the so-called sequential color method.

In discussing this color television system the following terms will be used:

color-field frequency—the highest vertical scanning frequency employed in the system;

color-frame frequency—color-repetition time per second, i.e., trichromatic-repetition rate per second, corresponding to the color-field frequency divided by three;

color-picture frequency—number per second of the coincidence of one and the same primary color with one and the same area of the image;

frame frequency—identical to the term used in monochromatic television, i.e., completion of the scanning of the entire picture area per second in black and white.

The system used employed sequential, additive-color scanning where the primary-color impulses of varying ratio, following in rapid succession, are integrated by the observer's eyes. The three primary colors employed were red, blue, and green, the characteristics of which will be discussed later. Rotating color disks or drums in front of the pick-up device and the receiving tube, suitably synchronized and phased, produced the color analysis at the transmitter and the synthesis at the receiver.

Various systems combining different interlace ratios, color fields, color-frame and color-picture frequencies as well as lines, were tried. If in a trichromatic system all areas of the image are to be scanned in all three primary colors, then the following conditions must be fulfilled:

$$c/f = 3n \pm 1,$$

where $c$ = color fields per second, $f$ = frames per second, and $n$ = any whole number 0, 1, 2, 3, etc. For sequential scanning $n=0$ and thus $c/f=1$, for $n=1$, $c/f$ becomes either 2 or 4, which corresponds to double and, respectively, quadruple interlacing. Figure 1 represents the values of the actual system arrived at as most advantageous

* For a more detailed account of color television, the reader is referred to a paper by the same authors appearing in the April, 1942 issue of *Proceedings of the Institute of Radio Engineers.*

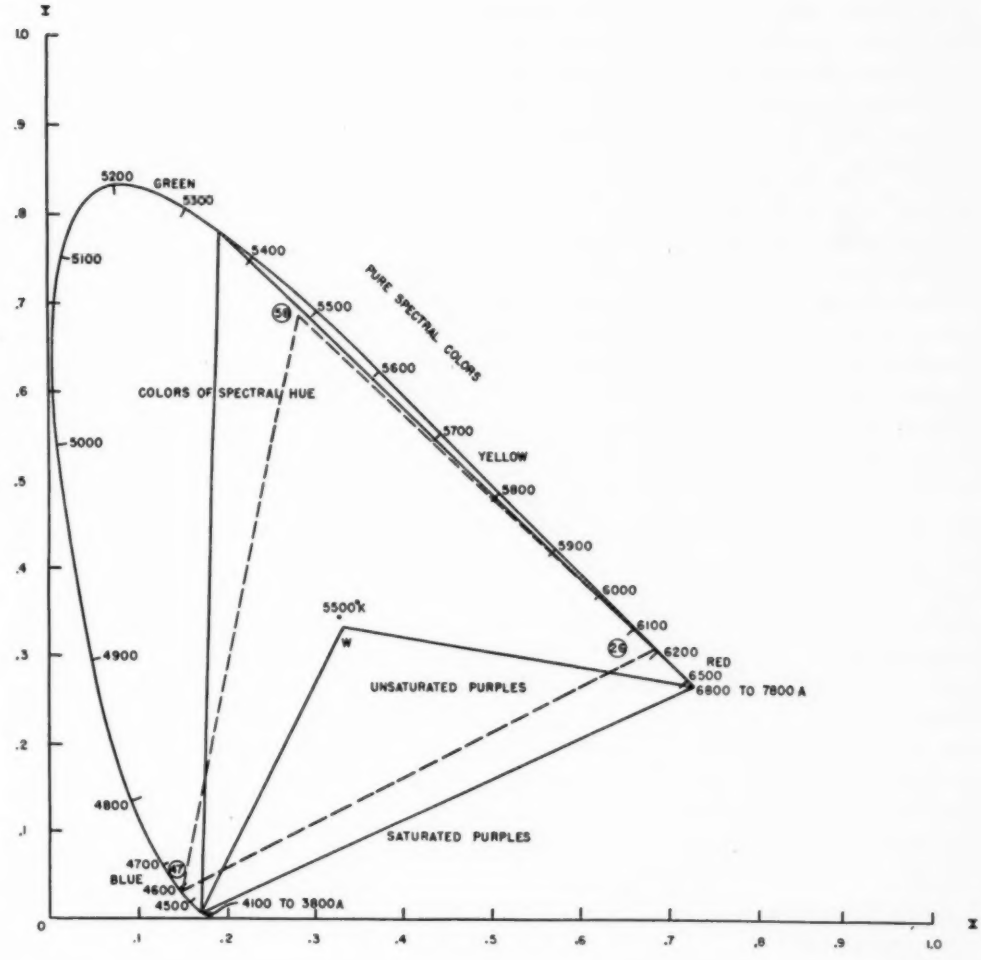



FIG. 2. Unified trichromatic coefficient diagram (color triangle).

within a 6-mc band and which was subsequently submitted to a field test.

## COLOR CHARACTERISTICS AT THE RECEIVER

The primaries at the receiver are determined by the color filters, red, green, and blue, and the fluorescent material in the tube. The largest gamut of colors is produced with primaries which fall on the locus of monochromatic colors in the color triangle. One such choice is red 7000, green 5350, and blue 4000 angstrom units, shown in Fig. 2.

Unfortunately, monochromatic primaries can only be obtained at the sacrifice of light intensity. Thus one finds that in television, as in certain color-reproducing processes, a compromise must be found between light intensity and the best choice of primaries. In addition, there is a restricted choice in available phosphors. The decay time of the fluorescent powder used in the receiving tube must be such that its intensity becomes negligible after one color-field period.

Of the commercially available phosphors the zinc and calcium sulphides possess sufficient luminescent efficiency and also satisfy the decay requirements. The luminescent spectrum of the phosphor must cover the entire range of the three filters in order to provide a light source for each primary. The precise character of the spectrum desired is contingent upon the choice of filters. The most desirable characteristic would be to furnish maximum light in those blue, green, and red regions that fit the maximum transmission portions of the blue, green, and red filters. Commercial tubes usable for color television employ, for the most part, two component mixtures utilizing a zinc sulphide with a spectral emission maximum in the blue and blue-green region, and

a zinc-cadmium sulphide with a maximum in the yellow and yellow-red region.[1] The spectral curve of such a mixture is shown in Fig. 3.

An inexpensive source of spectroscopically reproducible filters, with a wide selection of color, is the Wratten series, available in gelatin or acetate stock. The choice of the filters is determined by the wave-length at which the maximum transmission occurs, the width of the transmission band, and the total transmission. It is desirable to have filters and phosphor so chosen as to produce white corresponding to daylight fluorescent lamps with equal signals on the grid of the picture tube during the red, blue, and green periods. Thus if a white surface of 6000 to 6500 degrees Kelvin is transmitted, it should be received as the same shade of white and also should be identical to the receiver's own color when operated without a signal.

The filters finally chosen for use at the receiver were Wratten No. 26 for red, No. 47 for blue, and No. 58 for green. The emission curves for the phosphor mixture used for the experimental tubes combined with filters Nos. 47, 58, and 26 are given in Fig. 3. The resultant blue, green, and red primaries yield a new color triangle represented with broken lines in Fig. 2; the location of the new primaries is marked with the corresponding filter numbers.

A satisfactory method of specifying the color

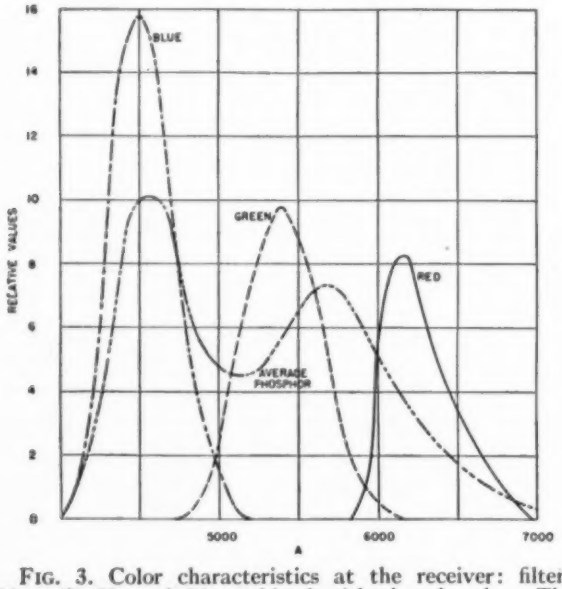



FIG. 3. Color characteristics at the receiver: filters Nos. 47, 58, and 26 combined with the phosphor. The fourth curve represents an average phosphor.

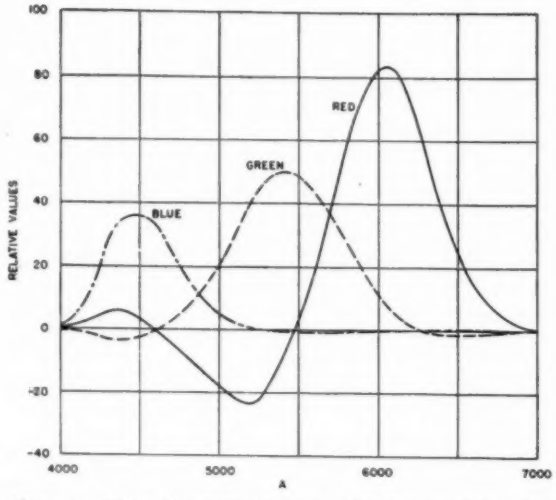



FIG. 4. Color characteristics at the transmitter: these curves are solutions of Eq. (9) on the basis of receiver characteristics represented in Fig. 3.

composition of the phosphor, which is a mixture of blue, yellow, and orange zinc and zinc-cadmium sulphides, without resorting to actual spectral curve data, was to specify the transmission through filters Nos. 47, 58, and 26 as recorded with a Weston photo-cell No. 2.

Figure 2 is a representation of conditions at the receiver based on the unified trichromatic coefficient diagram, sometimes called color triangle. The white produced with three equal signals appears somewhat bluish. Very recently, however, satisfactory "white" tubes were made in the laboratories, which show consistently good color characteristics and permit transmitter operation with equal blanking pulses. Spectral curves for these tubes are in preparation and will be available shortly for standardization purposes.

## TRANSMITTER COLOR CHARACTERISTICS

While the performance of the receiver was based on the color theory of vision, the study of the color characteristics at the transmitting end of the system had to be guided by the desirability of producing all colors encountered in nature. At the receiver three properly chosen narrow bands in the spectrum were sufficient; at the transmitter, the bands must be wide enough and sufficiently overlapping to produce a signal from every color. The exact character of the three spectral curves at the transmitter is determined by the filter and phosphor combination at the receiver.

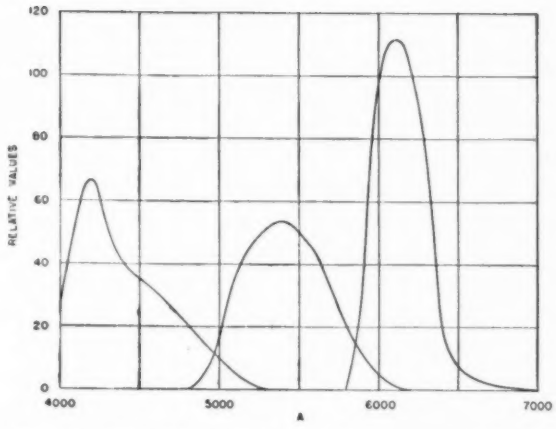



FIG. 5. Color characteristics at the transmitter: the curves combine the standard dissector, the carbon arc, 2 mm of Corning glass No. 978, and Wratten filters Nos. 25, 47, and 58.

The spectral characteristics at the transmitter are composed of the spectral sensitivity of the pick-up tube in the camera, the transmission curves of the filters, and the spectral emission of the light source. The filters and light source must be so chosen as to produce negligible signals in the infra-red (beyond 7000 angstrom units) and the near-ultraviolet (below 4000 angstrom units) regions of the spectrum.

The general relationship between the color characteristics of the transmitter and the receiver can be derived from an analysis developed by Hardy and Wurzburg in connection with photographic reproduction in color.[2] From calculations based on this analysis a transmitter color response has been arrived at and is shown in Fig. 4. The curves in this figure display the usual characteristics found in all color-reproduction problems, the existence of negative color values for perfect matching. The present color-television system has no mechanism to introduce these negative values; however, they are being partially compensated for with the aid of the color mixer.

## CAMERA TUBES

Before comparing the results contained in Fig. 4 with actual operating conditions, it is best to consider briefly the color characteristics of the tube used at the transmitting end. They are of two types: the dissector, used for slides and motion pictures, and the orthicon, used in the studio and in outdoor pick-ups. One of the problems in dissector operation is the elimination of signals produced with infra-red radiation. The infra-red contaminates all colors as it passes freely through the red, blue, and green filters. Originally the standard dissector was used with a cesiated silver-oxide cathode surface. This surface has a minimum color response in the green portion of the spectrum while it shows rising tendencies both towards the blue and the infra-red regions. In order to utilize this tube a carbon arc was used as a light source, combined with the infra-red-absorbing Corning glass filter No. 978, 2 millimeters thick. The filters that were used in this set-up were Wratten Nos. 47, 58, and 25. Figure 5 is a graphic representation of the results. The dissector used at present is the so-called daylight dissector (developed especially for color television by the Farnsworth Television and Radio Corporation) with a maximum in the green portion of the spectrum falling off towards the blue and the red end. This dissector was also used with a carbon arc but with a Corning filter No. 978 only 1 millimeter thick. The color filters were again Nos. 47, 58, and 25. In both cases a water cell was used to secure protection for the slides and film. Figure 6 gives the operating conditions for the daylight dissector. There is no question of the superiority of this type of tube over the standard dissector for this work. The signal-to-noise ratio is improved partly because of the greater photoelectric response in the pertinent portion of the spectrum and also because of the reduction in thickness of the infra-red filter.

The spectral characteristics of the orthicon

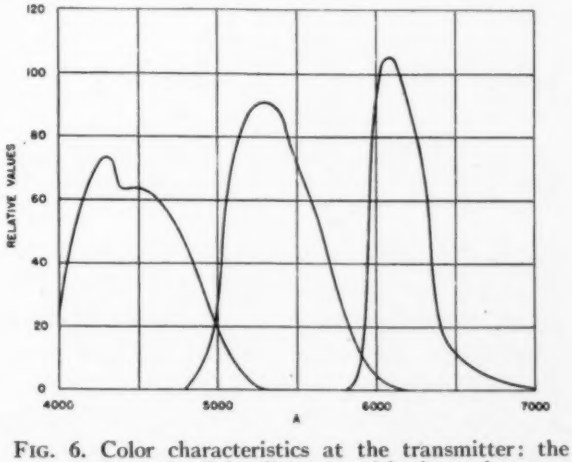



FIG. 6. Color characteristics at the transmitter: the curves combine the light dissector with the carbon arc, 1 mm of Corning glass No. 978, and Wratten filters Nos. 25, 47, and 58.

have not been measured under operating conditions. However, from measurements of the signal obtained through various Wratten filters it can be stated very generally that the spectral response of the color orthicon does not conform to the standard cesiated silver-oxide surface. It was found that specifications of the color characteristics of the orthicons used can be summed up in terms of the signal obtained through Wratten filters Nos. 47, 57, and 25. The ratio of the signals through No. 47 to No. 57 using daylight fluorescent light source is approximately 1.25 and the ratio of No. 47 to No. 25 is 0.85. To obtain a general view of the color characteristics at the transmitter a daylight fluorescent-light source attenuated through filters Nos. 47, 57, and 25 has been plotted in Fig. 7. The inclusion of the characteristics of the orthicon in this figure will reduce the red curve roughly by 30 percent. It may shift the maxima slightly but it will not change the limits of the individual curves. Thus the extent of filter overlapping remains the same.

## FLICKER IN COLOR TELEVISION

The well-known Ferry-Porter[3, 4] law states that the critical frequency is proportional to the logarithm of the illumination intensity. Porter was the first to establish the fact that the critical

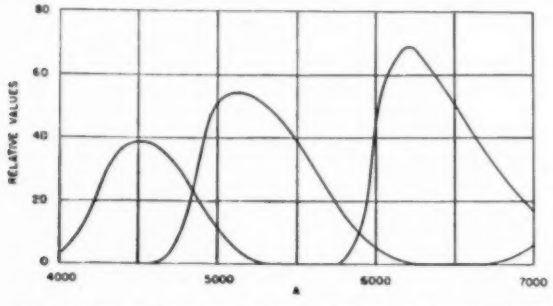



FIG. 7. Color characteristics at the transmitter: Wratten filters Nos. 24, 47, and 57 combined with daylight fluorescent light source as used with orthicons.

fusion frequency is independent of wave-length, provided the apparent brilliance remains constant. Based on this law he derived the well-known color-intensity curve of the eye using a flicker photometer. It follows that with constant irradiation over the visual spectrum the eye's sensitivity to flicker follows the color intensity curve.

In the sequential color-television system under consideration the worst flicker condition would occur if only the primary color with the highest luminosity were received while the other two primary colors were suppressed completely. Such a condition hardly ever occurs in practice unless the color camera picks up a green object, the limits of its chromaticity being between 5400 and 5800 angstrom units (see Fig. 5). The flicker frequency for this case would be 40 per second. An attempt will be made to calculate the maximum permissible brilliance of such a green picture before flicker becomes perceptible.

The validity of Talbot's law has been checked for all colors by Porter[4] and others. Thus, applying Talbot's law to the special case of color television as discussed here, the apparent brilliance of the image at the receiver would be

$$I_A=\frac{1}{T_c}\int_0^T L(t)dt, \qquad (1)$$

where $T_c$ is the duration of a complete color cycle (color frame) and $T$ the duration of a color field. $L(t)$ is the decay function of the screen material. This is assumed to be exponential, with a luminosity not greater than $\frac{1}{10}$ of the initial brilliance after the duration of one field period ($T$).

Thus for the transmission of green between 5400 and 5800 angstrom units the apparent brilliance of the received picture becomes

$$I_g=\frac{1}{T_c}\int_0^T Y_g L(t)dt, \qquad (2)$$

where the luminosity at the receiver has been expressed in terms of the $Y$ component of the tristimulus coefficients, representing the combination of the receiver's green filter and the screen material, as shown in Fig. 2, and the decay with time of the screen material.

For a picture tube made in the laboratories especially for color reception the following ratios of the luminosity values of the three color primaries were found:

$$\frac{Y_{\text{green}}}{Y_{\text{blue}}}=23 \quad \text{and} \quad \frac{Y_{\text{red}}}{Y_{\text{blue}}}=10.3. \qquad (3)$$

Equation (2) can be rewritten

$$I_g=\frac{Y_g}{T_c}\int_0^T L(t)dt.$$

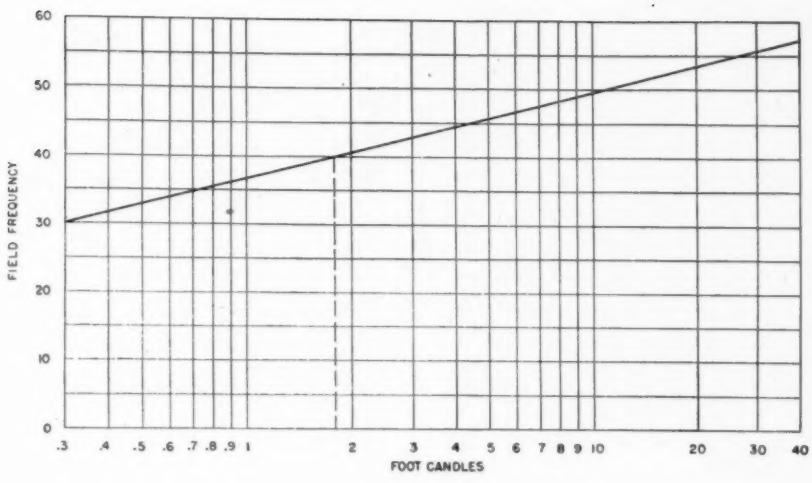



FIG. 8. Relationship between critical flicker frequency and picture brilliance.

Using equation (3) we obtain

$$I_g=\frac{{}_{23}Y_b}{T_c}\int_0^T L(t)dt. \tag{4}$$

In order to determine at what apparent brilliance a 40-cycle television picture will just begin to show flicker we consult the curve shown in Fig. 8.[5] Accordingly a repetition frequency of 40 per second will permit a screen illumination of 1.8 apparent footcandles.

So far we have considered the most unfavorable case from a flicker point of view. More favorable conditions will occur if white with three equal electrical impulses during the red, green, and blue periods is transmitted and received.

In order to obtain the total apparent illumination at the receiver, (2) will be expanded into

$$I_w=\frac{I}{T_c}\int_0^T Y_gL(t)dt+\int_0^T Y_rL(t)dt+\int_0^T Y_bL(t)dt; \tag{5}$$

assuming that $L(t)$ is independent of color, Eq. (5) can be written as

$$I_w=\frac{(Y_g+Y_r+Y_b)}{T_c}\int_0^T L(t)dt. \tag{6}$$

Using the substitutions in (3) we obtain

$$I_w=\frac{34.3Y_b}{T_c}\int_0^T L(t)dt. \tag{7}$$

Comparing this with (4) it is found that

$$\frac{I_w}{I_g}=\frac{34.3}{23}. \tag{8}$$

The meaning of this expression is that the apparent brilliance of the receiver screen increases only in the ratio of 34.3/23 when white is transmitted, even though it is produced by three equal electrical impulses, one during each color field. Since in this case a light impulse is received during each of the color fields, flicker conditions improve rapidly. However, the apparent brilliance should not be higher than 1.8 times 34.3/23, which is 2.7 apparent footcandles if one wishes to make sure that in the singular case of the transmission of a narrow band of green no flicker is present. This top value of 2.7 footcandles is not a serious limitation. Present black-and-white pictures with such a high-light brilliance give satisfactory viewing in darkness. The same black-and-white receiver, however, will not permit satisfactory viewing with surrounding illumination of any appreciable magnitude.

Color-television pictures produced with the aid of rotating filters do not deteriorate appreciably in the surrounding illumination due to the fact that the room light which passes through the filters twice is attenuated by the square of the filter loss factor, while the picture itself is only attenuated by the first power of the filter factor.*

* A black-and-white receiver with an equivalent neutral filter will resist room illumination to the same extent.

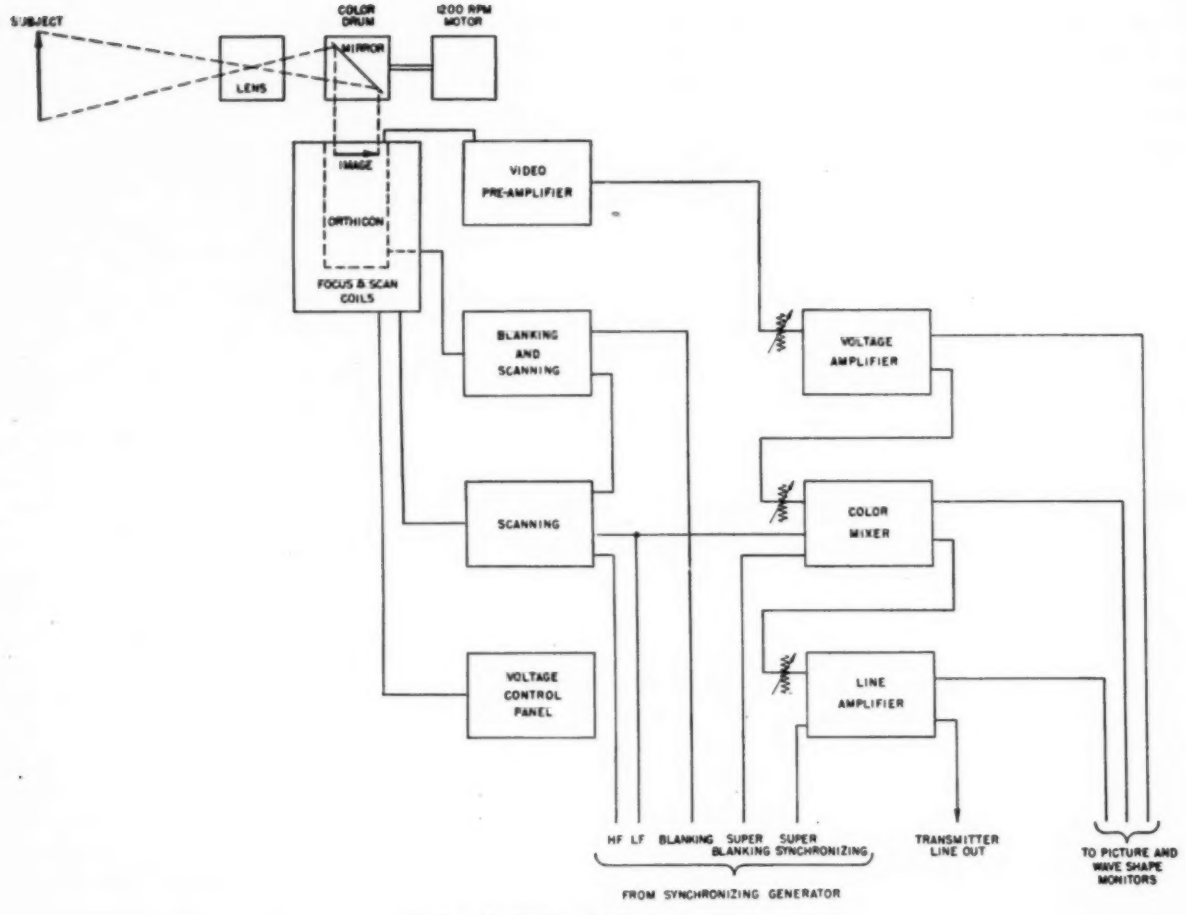



FIG. 9. Orthicon block diagram.

As a result the unmodulated screen of a color receiver appears nearly black even in a well-illuminated room. Thus the 2.7-footcandle maximum brilliance will furnish a more satisfactory image than a conventional black-and-white receiver even with 10-footcandle high-light illumination.

## EQUIPMENT FOR COLOR TELEVISION

### Studio

Certain electrical requirements must be met by pick-up tubes if they are to be used in color television. It is important that the signals produced during any one color field should not be adulterated by a signal left over from a previous field. Storage-type camera tubes must, therefore, be designed so that the entire electrical charge on the mosaic is removed within one field period.

A constant black level must be established in the camera tube, and spurious signals such as "shading" should be absent. The dissector is the only commercial camera tube that meets all of the above requirements though its usefulness is limited to the transmission of film or slides due to its low sensitivity.

The orthicon as modified for color television with lower mosaic capacitance was developed through the cooperation of the RCA Radiotron Division and has been found to produce very acceptable color pictures with incident light of 150 footcandles on the subject. A certain amount of "hangover," which may be defined as the amount of signal remaining on the mosaic after the scanning beam has completed one field, appears to be unavoidable, but is only troublesome at lower light levels. The hangover occurs when the potentials on the mosaic are small enough to be within range of the random velocity distribution of the scanning beam. Under this condition the charge is not completely removed

until the mosaic has been scanned several times. A lower mosaic capacitance permits the voltage on the mosaic (for a given illumination) to build up to higher levels.

It might be expected that difficulty would be experienced with interlaced scanning on a storage-type camera tube such as the orthicon. In an ideal storage tube, where the scanning spot is only one line wide, it will be apparent that at the end of one field-scanning period only one-half the lines will be scanned and the hangover will be 100 percent. Actually, in practice, hangover ratios of from 1 : 5 to 1 : 10 are obtained with the orthicon, indicating that either leakage, defocusing, or other effects are present to such an extent as to remove most of the unscanned picture at the end of one field scansion.

The gamma of a camera tube need not necessarily be unity, as correction may be made for any particular characteristic later on, if desired. In general, a television system employing a linear pick-up tube such as an orthicon, will have an over-all gamma higher than unity, due to the cathode-ray tube. A reduction in the gamma may be more satisfactorily made with tubes of the dissector type, where the noise is negligible in the black portions of the picture, than in a tube of the orthicon type, where the noise is determined by the impedance of the tube and the first amplifier stage. A reduction of gamma in the latter case results in increased noise in the blacks, if all other factors remain equal.

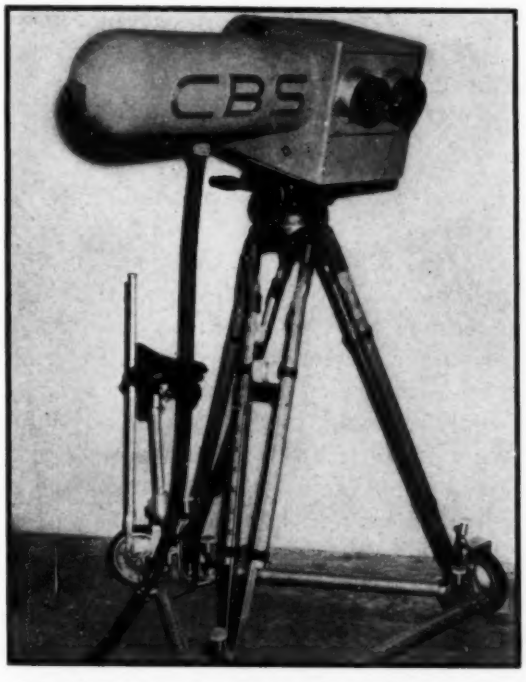



FIG. 10. Orthicon camera on tripod; direct-pick-up color camera.

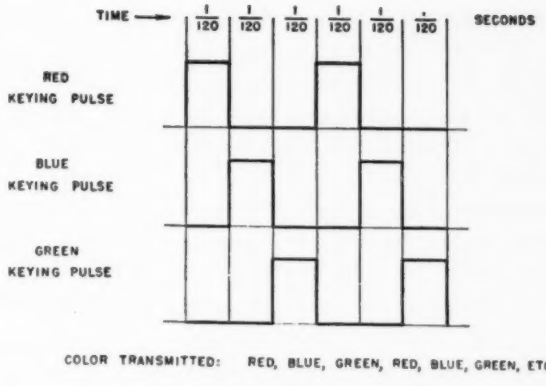



FIG. 11. Color mixer pulse diagram.

The introduction of color does not change many of the design requirements which are ordinarily encountered in monochromatic television studio equipment. Certain factors, however, are worthy of mention. The color-field frequency of 120 per second necessitates freedom from 60-cycle hum in the synchronizing generator and scanning equipment and, to a lesser degree, in video amplifiers. Sixty-cycle components present in the synchronizing generator or scanning equipment cause loss of interlace and in the video equipment cause flicker at a 20-cycle rate resulting from the beat between the 60-cycle hum and the 40-cycle picture components. Hum may be eliminated easily by operating the equipment from a 120-cycle power source.

Good low frequency response is necessary in video amplifiers to pass the 40-cycle picture components properly. The video control equipment for color is somewhat more complex than for black-and-white transmission, as it seems advantageous to control the gain and possibly the background of each color independently, as previously mentioned.

Block diagrams of a color-television system using an orthicon camera tube are shown in Fig. 9. A photograph of the orthicon direct-pick-up color camera is shown in Fig. 10. As previously mentioned, it is essential that the black level be established at the camera since manual control of the direct-current level for each color would

seem a tremendous task. This is done by applying the blanking pulses to the grid of the orthicon and to the cathode of the dissector.

The video amplifiers in a color-television camera channel are conventional except for the color mixer. The color-mixer amplifier may be described as an electronic switch combined with three separate amplifiers, each with its own gain and brightness controls. The video signal is switched by means of suitable timing pulses (Fig. 11). The pulses are so timed as to operate each of 3 amplifiers in succession, turning on one as another is turned off. This switching occurs during the blanking period, and switching transients are removed by subsequent clipping of the recombined signals. Blanking is injected on the cathode of each switching amplifier tube and the individual brightness of each color is adjustable by bias controls.

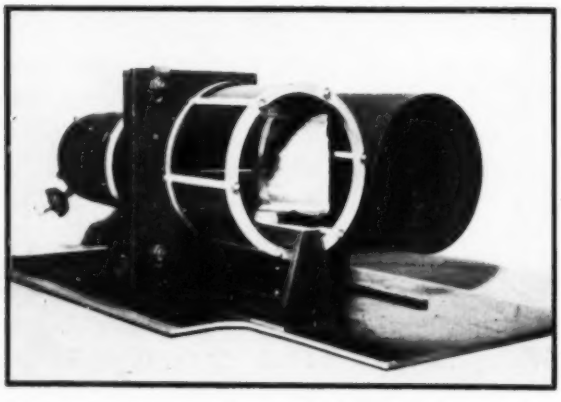


FIG. 12. Orthicon color camera; filter drum assembly.

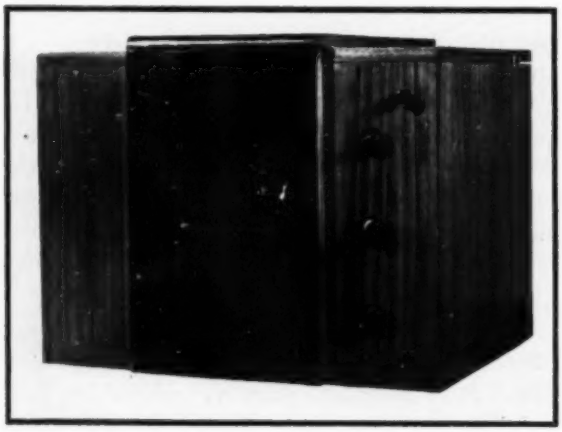


FIG. 13. Seven-inch color television receiver; front view.

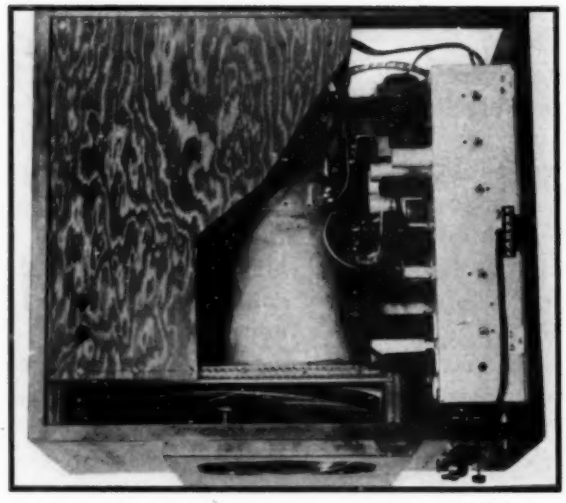


FIG. 14. Seven-inch color television receiver; top view, open.

The non-storage dissector is more easily adapted to a filter disk, since it is necessary that the optical image on the cathode be of the correct color only at the point which is being scanned. The orthicon on the other hand, being of the storage type, requires that only one color can be present in the optical image for one complete scanning-field period, prior to the actual scansion of a given point. The filter drum which provides color selection at the pick-up tube is shown in Fig. 12. It occupies less space than a color disk of corresponding characteristics.

## RECEIVING EQUIPMENT

Practically any good black-and-white television receiver design may be made the basis of a color-television receiver. A typical color receiver is shown in Fig. 13. Figure 14 shows the same receiver open, looking in from the top. The additional equipment required will be the color disk, with its driving and synchronizing means.

Usually additional precautions should be taken to insure complete direct-current component insertion and freedom from hum. The cathode-ray tube should be magnetically shielded to minimize pairing due to 60-cycle hum. It is fully as important, however, that attention be paid to sources of hum in the scanning circuit, such as common ground returns carrying alternating current, insufficient power-supply filtering, or magnetic coupling from power equipment into

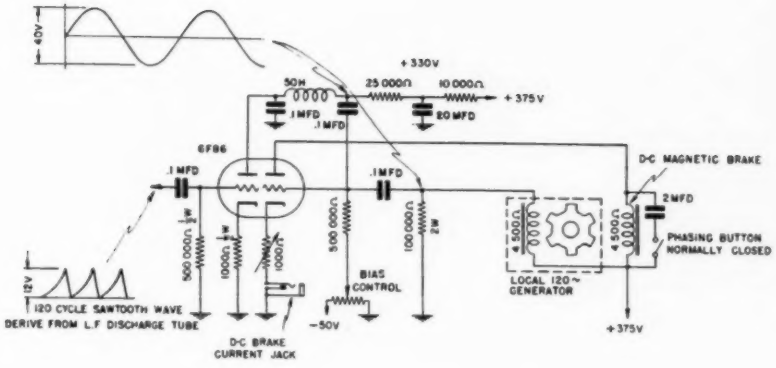
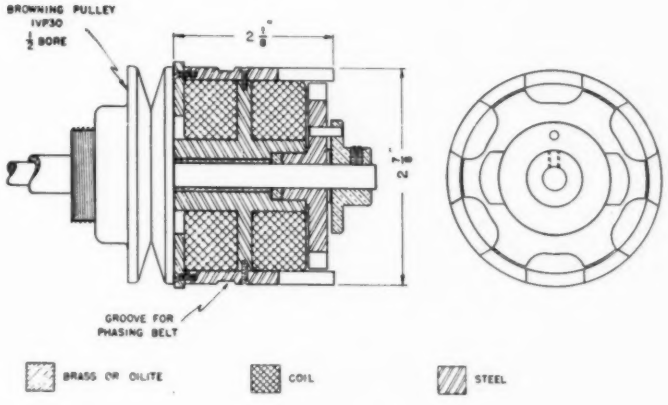



FIG. 15. Seven-inch receiver brake circuits and design.




scanning transformers. It has sometimes been found economical to inject hum in opposing phase to neutralize alternating-current fields that otherwise may be difficult to remove. This is only possible when the interfering hum remains constant in phase and amplitude with respect to the neutralizing components. Another effect which sometimes tends to destroy interlacing is due to electrostatic charges which accumulate on the rapidly moving color disk. Variations in the charge over the surface of the disk produce movement of the scanning lines as the disk rotates. It is possible to remove the charge with a semiconductive coating on the cathode-ray-tube face or with other electrostatic means of screening.

Color disks have been made of metal or of transparent plastics such as Lucite, Plexiglass, etc. Wratten filters may be obtained coated on a 10/1000-inch acetate stock which can be riveted to the plastic or metal disk. The disk may be rotated by a synchronous motor or by an asynchronous motor with auxiliary synchronizing means. Owing to lack of synchronism between the power supplies of New York, Connecticut, and New Jersey, and also owing to lack of standard synchronous motors of 1200-revolution-per-minute type, it was found desirable to drive the disk with an inexpensive induction-type motor and synchronize it by means of a phonic motor or a magnetic brake. Satisfactory phonic motors have been constructed which are driven by a single 6V6 tube, but the brake arrangement

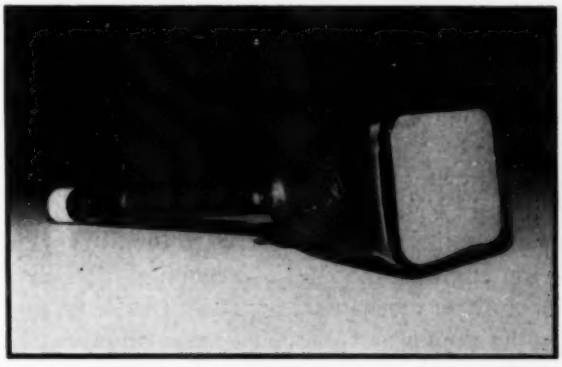


FIG. 16. Rectangular flat-screen cathode-ray tube for color television.

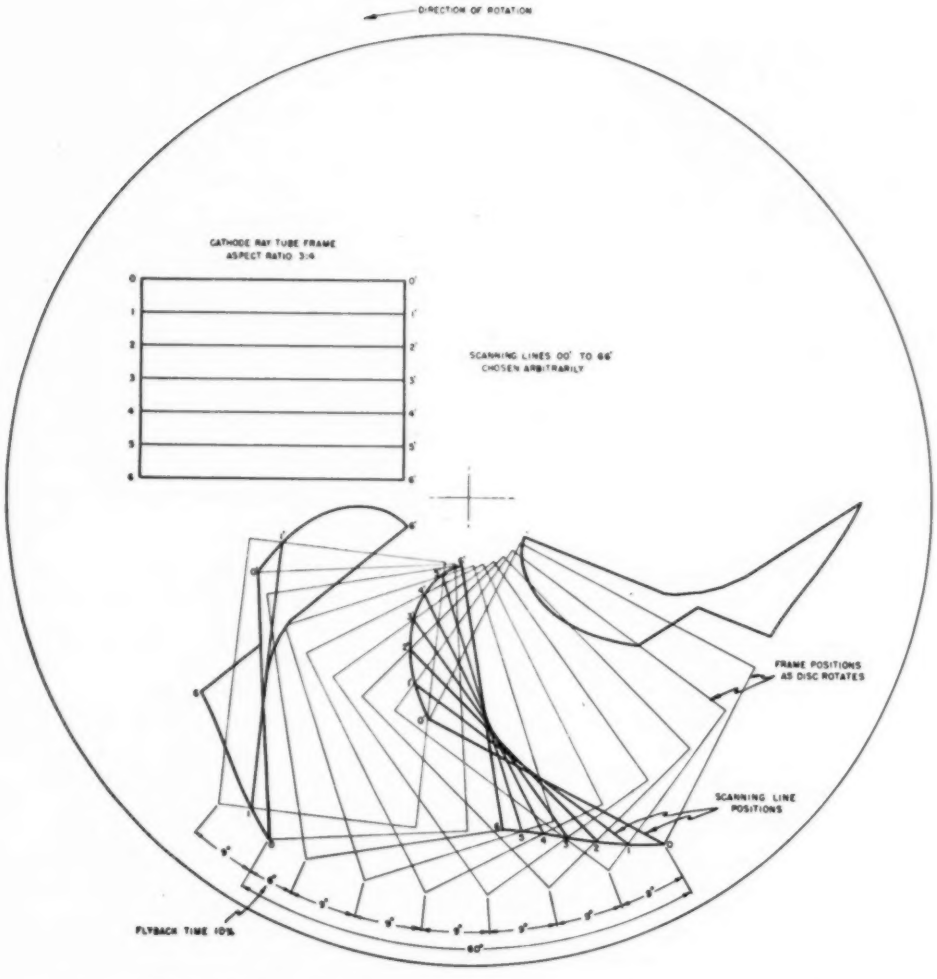



FIG. 17. Generation of filter shapes.

is preferred, Fig. 15. The 120-cycle voltage is derived from the low frequency scanning circuit and is mixed with a similar voltage from a small generator on the disk shaft. The sum of the voltages is then rectified and the resulting direct current applied to the magnetic brake. A departure in the disk phase with respect to the scanning produces a corresponding correction on the part of the brake.

The generation of a properly shaped filter disk is shown in Fig. 17. This type of shape is suitable for a receiving or transmitting tube with short decay or storage times. The curve which is obtained in Fig. 17 is an envelope of the position of a scanning line as traced onto the filter which is moving with the line. The required filter shape for a given mechanical arrangement is obtained by developing curves which make allowance for positive and negative tolerances to take care of fluctuation in the disk position, viewing angle, and screen decay, Fig. 18. Generally, the minimum disk diameter is about twice the outside diameter of the tube plus 1 or 2 inches. The optimum location will be determined by such factors as the distance from the disk shaft to the picture frame, but can be determined for any particular arrangement.

Color drums have been used at the receiver as well as at the transmitter instead of color filters. A short cathode-ray tube can be placed within the drum. The drum is designed for a lower speed of revolution than is usually possible with the disk. Successful drums have been built to operate at a speed of 600 revolutions per minute or one-half the usual disk speed.

The small table-model receiver shown in Fig. 13 utilizes a cathode-ray tube developed especially for color pictures. The screen of the tube is flat and has the exact shape of the final image (Fig. 16). The tube produces a picture equivalent in

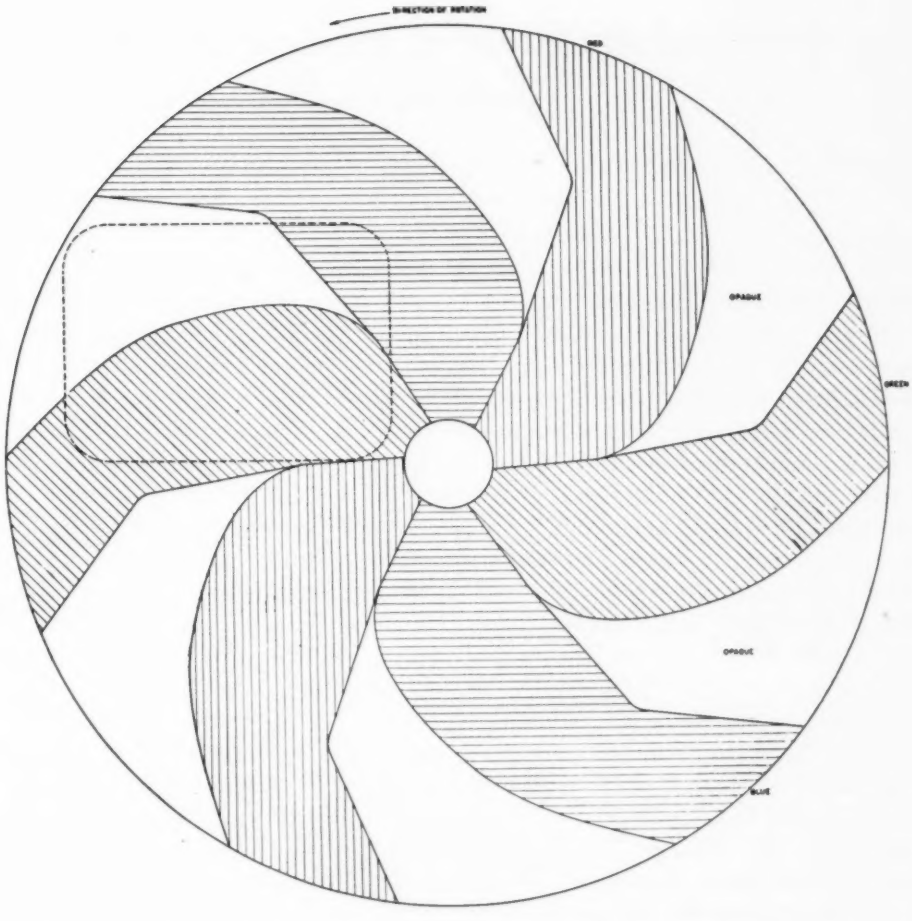


FIG. 18. Typical receiver filter disk design.

size to that of a conventional 7-inch round tube. Thus the color disk is 15 inches in diameter. A 10-inch lens with a focal length of 12 inches, which is built into the receiver, increases the image to correspond to that of a conventional 9-inch tube. Owing to low magnification, distortion and decrease of the viewing angle are appreciably reduced.

A history of color television, a bibliography and more detailed mathematical analysis of color characteristics together with circuit diagrams are contained in an article on this subject in the April, 1942 issue of the *Proceedings of the Institute of Radio Engineers*, and have been omitted here in order to preserve space.


## ACKNOWLEDGMENT

All members of the Columbia Broadcasting System television engineering department have been concerned with this development. In particular, the active cooperation of Messrs. Goetz, Freundlich, Harcher, Erde, Doncaster, Reeves, and Haas is gratefully acknowledged.

The CBS television program department has unselfishly contributed by preparing suitable test and demonstration material.

Thanks are due to Mr. Adrian Murphy for his constructive criticism, active encouragement, and enthusiastic support throughout the entire developmental period.


## BIBLIOGRAPHY


(1) H. W. Leverenz and F. Seitz, "Luminescent materials," J. App. Phys. **10**, 479–493 (1939).

(2) A. C. Hardy and F. C. Wurzburg, Jr., "The theory of three-color reproduction." J. Opt. Soc. Am. **27**, 227–240 (1937).

(3) E. S. Ferry, Am. J. Sci. **44**, 192 (1892).

(4) T. C. Porter, "Study of 'flicker,' " Proc. Roy. Soc. (London) **70**, 313–329 (July 29, 1902).

(5) E. W. Engstrom, "A study of television image characteristics," Proc. I. R. E. **23**, 295–310 (April, 1935).

# Physical Methods of Dosage Determination in Radiation Therapy

BY EDITH H. QUIMBY

*Memorial Hospital, New York, New York*

IN the determination of dosage in x-ray therapy, two types of information are desired, a measurement of the quantity of radiation and a specification of its quality. The purpose of the present paper is to discuss the measurement of quantity. Since, however, the methods employed for this purpose depend to some extent on the quality, it is desirable to consider this briefly.

The qualitative attribute of radiation with which one is concerned in radiation therapy is its penetrating power. This is dependent on the wave-length, which is dependent on the exciting voltage and other factors. The higher the voltage the shorter the wave-length of some of the components of the beam, and the shorter the wave-length the more penetrating the x-rays. Any beam used in therapy is composed of a heterogeneous mixture of wave-lengths, the minimum being defined by the voltage, but the total composition being dependent on the voltage wave form, the filter, etc. Hence no single statement of voltage or wave-length describes it accurately. Various schemes for defining an "average" wave-length, or an "effective" wave-length have been proposed; none has been generally accepted. Since the penetrating power is the factor about which information is desired, a method of quality specification referring directly to that is now generally employed. It is the statement of the "half-value layer." This is the amount of a specified material which reduces the intensity of the beam by 50 percent. Thus the more penetrating the radiation the higher the half value layer. In general, within the range of x-rays used in therapy, the half-value layer is specified in aluminum for voltages up to about 140 kv, in copper from 140 to 500 kv, and in lead from 500 kv up. Whenever in the course of this paper the quality of the beam is mentioned, the penetrating power is to be understood.

In the determination of quantity of radiation, or "dose," the problem is twofold. On the one hand, it is necessary to have some method of calibration, to insure a knowledge of the output of the x-ray tube and its variation with changes in any of the physical factors. On the other, it is desirable to know the amount of energy actually absorbed in any specified region within the body of the patient, in order to plan the treatment properly. While these can be handled as separate problems, it is, of course, preferable to coordinate them as far as is practicable.

For calibration of the tube, many possibilities are open, while in the measurement of dosage to the patient certain very definite limitations arise. These limitations should be considered before a decision is made as to the method for calibrating.[1,2]

Any reaction produced by radiation is due to the energy absorbed from the beam, and of course from this point of view the actual dose delivered to any biological organism is the energy absorbed by it. This it is not possible to measure directly at the present time. The energy absorbed in the biological medium produces ionization therein; it is this ionization which is believed to be largely responsible for the biological effects produced. Not all of the absorbed energy manifests itself as ionization, but, so far as is known, the ratio between the total absorbed energy and that utilized in forming ions is essentially independent of the wave-length of the x-rays, over the range ordinarily employed in therapy. Hence a measurement of the ionization produced within the biological medium would give something proportional to the energy absorbed therein. This cannot be made either. However the energy absorbed per gram of tissue is the same as that which would be absorbed per gram of gas having the same chemical constituents in the same proportions, under the same conditions. This absorption depends on the atomic numbers of the components of the absorber, and since in living tissues these elements are all of low atomic number, the net effect does not differ much from that

of air, within the wave-length range under consideration. Hence one may reasonably expect that ionization in air will have a constant proportionality to energy absorbed in tissue, provided the measurements are made under suitable conditions. For these reasons, ionization in air has been adopted as the general basis for measuring systems, both for the calibration of x-ray equipment, and for the specification of something which is called the *dose*, and which is actually proportional to the energy absorbed at the location for which the dose is specified.

### A. CALIBRATION OF X-RAY OUTPUT

For purposes of calibration, the ionization produced in air at a certain position in the beam can readily be measured by the use of a simple electroscope or electroscope and ionization chamber. By this means, *relative* values of x-ray outputs for different machines or different physical factors can be determined by carrying a single instrument, or identical instruments, from place to place. If a particular instrument could be agreed upon as standard, such a procedure could be generally adopted; the unit of output would then be something associated with a particular instrument.

This, of course, is unsatisfactory; a unit which is not a function of any special instrument is desirable, with some sort of standard device for establishing the unit. This device need not be a simple and portable instrument; however if it is not, it is necessary that simple and portable secondary instruments can be calibrated against it.

The unit which naturally suggests itself, in the light of the energy considerations just discussed, is that quantity of radiation which will produce some specified number of ion pairs in a unit volume of dry air under standard conditions. Such a unit was essentially agreed upon by radiological physicists almost twenty years ago, and everyone was glad to name it the *roentgen*. However for a long time there were controversies among these physicists as to certain details of the definition of the unit and of the physical measurements to be made for the purpose of establishing it.

Finally, at the Second International Congress of Radiology, in 1928, the following definition was adopted: "The unit of dose is that quantity of roentgen radiation which, when the secondary electrons are fully utilized and the wall effect of the chamber is avoided, produces in 1 cc of atmospheric air, at 0°C and 760 mm of mercury pressure, such a degree of conductivity that 1 electrostatic unit of charge is measured under saturation conditions. This unit shall be called the roentgen, and designated by *r*." At the Fifth International Congress of Radiology, in 1937, the definition was modified as follows: "The roentgen shall be that quantity of x- or gamma-radiation

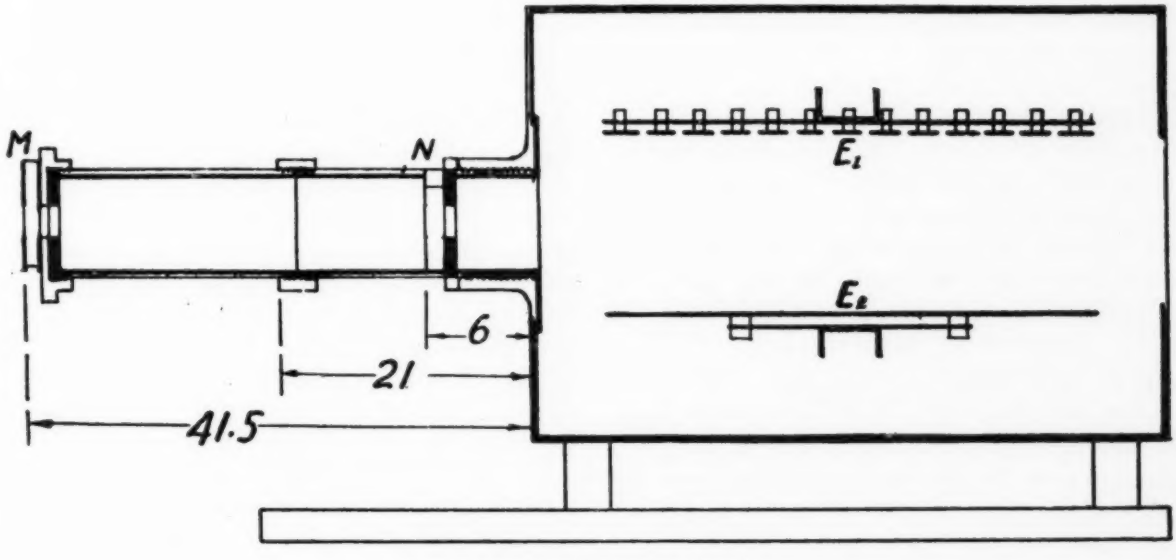



FIG. 1. Standard ionization chamber.

such that the associated corpuscular emission per 0.001293 gram* of air produces, in air, ions carrying 1 e.s.u. of quantity of electricity of either sign."[3]

Some significant points about these specifications should be emphasized. The secondary electrons must be fully utilized in producing ions, hence they must not strike the wall of the ionization chamber before their energy is completely expended. There must be no wall effect, that is, neither the primary x-rays nor the secondary electrons must strike any scattering material which can introduce extraneous secondary radiation. It is evident that for the strict fulfilment of these requirements an open air ionization chamber of large dimensions is necessary. For this purpose the so-called "standard ionization chamber" has been developed; one design of such a chamber is shown diagrammatically in Fig. 1.[4] In this particular chamber the outside case is $36\times36\times53$ cm and is made of brass with a 3-mm lead lining as protection from scattered radiation. The electrodes $E_1$ and $E_2$ are parallel plates of aluminum $22\times42$ cm. $E_1$ is a single plate, and is charged to the necessary potential to provide saturation current, i.e., to remove all of the ions of one sign from the region between the plates before any of them can recombine. This potential in any particular case depends on the number of ions per cc per second and the distance between the plates. $E_2$ is made up of eleven parallel strips set very close together. Any strip can be either connected to the electrical measuring instrument, or grounded. In practice one or a few of the central strips are connected to the measuring instrument and form the collecting electrode; the others are grounded, thus furnishing effective guard plates for the collector. The distance between the electrodes can be varied; it must always be great enough so that the energy of the secondary electrons is completely used up within it. The distance to be used is therefore dependent on the quality of the primary beam.

The x-ray beam entering the chamber is defined by a diaphragm system which must satisfy certain conditions depending on the size of the focal spot, the distance from the target to the chamber, etc.

The information to be obtained with this chamber is the number of ions produced per cc of air at a specified position. This necessitates an accurate knowledge of the ionized volume, as well as an accurate measurement of the ionization current. It has been shown theoretically by Behnken, Taylor, and others that in such a chamber the ionized volume is to be considered as the product of the area of the limiting diaphragm and the dimension of the collecting electrode parallel to the axis of the beam; the location to which the ionization per cc thus obtained is to be assigned is the position of the limiting diaphragm.[5,6] For the actual measurement in e.s.u. of the very small currents involved, various methods have been developed. Problems of measurement of such small currents are not peculiar to x-ray dosage; it is not desirable to include a discussion of them in the present paper.

By the use of such a standard ionization chamber, the number of roentgens per unit time at any specified place in air can be determined. The instrument, however, is not satisfactory for calibrations, for it cannot be readily made in convenient portable form, and its use demands more precautions than are generally employed in the ordinary hospital x-ray department. However, small convenient closed ionization chambers can readily be calibrated against the standard open air instrument, and used as secondary standards for determining x-ray output wherever it may be desired.

When a closed ionization chamber is employed instead of one of the open air type, it appears that the ionization which is being measured in the two cases is not the same. The wall of the (small) closed chamber absorbs some of the primary beam, sets up its own secondary radiation (the wall effect referred to in the first of the above definitions) and does not allow space for the complete utilization of the secondaries. Thus the relation between the readings with the two chambers might be expected to be different for every quality of primary radiation. This is indeed the case with secondary chambers made of metal. Such are obviously of little value for calibrating.

However, it has been shown by Fricke and Glasser, Victoreen, and others that when a small closed chamber is composed of organic material, of approximately unit density, with an effective

* The mass of 1 cc of dry air under standard conditions.

atomic number essentially the same as air, constructed according to certain geometrical and other conditions, the ratio of its reading to that of the standard chamber is constant over a wide range of quality.[7] Such chambers are said to be "air-wall"; it appears that within this range of constant relationship the ionization per cc in the chamber is the same as it would be at the same point in the beam in free air.

For the ordinary range of therapeutic x-rays, from about 100 to 250 kv, such chambers are satisfactory as secondary standards. For radiation beyond these limits, modifications must be made. In the case of "harder" radiation (short wave-length, great penetration) the chamber walls must be made thicker than for the usual qualities. This is because for the correct measurement of x-rays in roentgens, the primary beam must be in equilibrium with its secondaries; this is implicit in the condition of complete utilization of secondaries. For this situation to exist, the chamber wall must have a certain thickness. For x-rays up to the region of 200 kv, the required thickness is very small—a fraction of a mm. However as the voltage increases and the range of the secondary electrons becomes greater, more material is necessary. For million volt x-rays, 2-mm walls are necessary; for the gamma-rays from naturally radioactive substances the required thickness is about 4 mm.[8] At the other extreme, for very "soft" radiation (long wave-length or little penetrating power) the absorption in even quite thin walls becomes appreciable, and the ionization per cc in the closed chamber becomes less than in the air.

In the region of 45 kv another factor enters, namely, the effect of photoelectrons from the walls of the chamber.† The inner walls of these small chambers are coated with India ink, Aquadag, or other carbon preparation, to make them electrically conducting without introducing elements of atomic number higher than those in the walls. The contribution of carbon photoelectrons to the ionization in the chamber is not significant unless the beam contains x-rays of very long wave-length; this is the case for the 45-kv radiation. Since the atomic number of carbon is less than the effective atomic number of air, the result is to produce less ionization than if all the secondaries were from air. This effect is more noticeable the narrower the air space in the ionization chamber. Therefore, in this quality range very small** closed chambers are not satisfactory for calibration, because of both absorption of primary radiation in the walls and the effect of the photoelectrons from the carbon coating.

With these secondary calibrating chambers, various devices are employed to measure the electric current. They may be of the same sort as are used with the standard chambers; in this case (as with the standard chambers) certain precautions must be taken regarding the electrical leads and connections, and provision must be made to insure the protection of the measuring instrument and the operator from primary and scattered x-rays. More general in this country at the present time is the use of the so-called "condenser dosimeter," usually the Victoreen instrument, equipped with the Glasser-Seitz condenser chamber. This warrants some description because of its rather widespread adoption in nuclear physics laboratories for the measurement of neutrons, in addition to its original function in measuring x-rays. It is shown in Fig. 2.

The thimble-shaped ionization chamber is connected to a condenser, which is in turn connected to the electrometer. The tube containing ionization chamber and condenser can be disconnected from the box containing the electrometer. In practice the procedure for the use of this instrument is as follows: The chamber and condenser section is attached to the electrometer, and the whole system charged, by means of a static charger inside the box, until the electrometer fiber indicates a certain position on the scale. The removable section is then detached, placed so that the ionization chamber is at the desired position in the x-ray beam, and exposed for a meas-

† This is not to be interpreted as a critical voltage at which this phenomenon appears. It is the operating voltage for a type of x-ray equipment clinically employed in the treatment of very superficial lesions. The tube has very small inherent filter, so that the emitted radiation is very soft (half-value layer 0.3 mm Al). For somewhat higher voltages with equally small filtration no investigation has as yet been made.

** Very small chambers are necessary in practice with the type of x-ray tube used at this voltage, because of (a) the very high intensities (of the order of 10,000 r per minute), and (b) the very short target-skin distances employed.

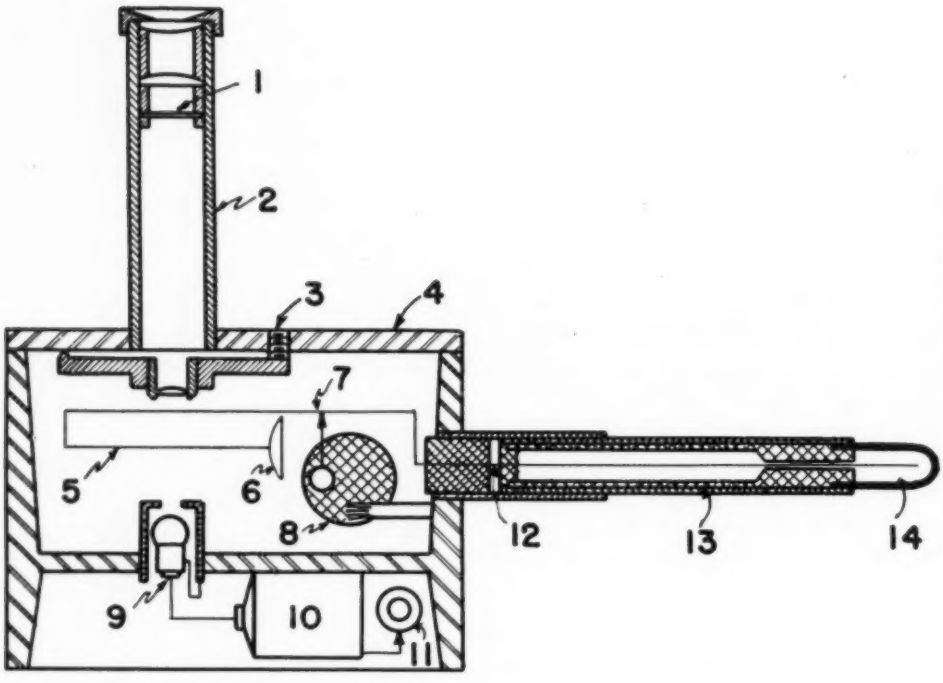



FIG. 2. Victoreen measuring instrument. (1) Scale to read deflection of string, (2) microscope, (3) focus adjustment, (4) lead case, (5) platinum string, (6) quartz loop, (7) switch, (8) amber charging disk, (9) lamp, (10) batteries, (11) lamp switch, (12) contact, (13) condenser (amber), (14) thimble ionization chamber tube. (Courtesy of Victoreen Instrument Company.)

ured time interval. The ionization in the chamber partly discharges the condenser. This section is then re-attached to the electrometer, the drop in charge of the condenser causes the fiber to move to a new position on the scale. The scale is calibrated in roentgens at the factory, hence with a knowledge of the exposure time and of the roentgens recorded, the roentgens per minute at the point of measurement are immediately known.

### B. DETERMINATION OF DOSAGE TO THE PATIENT

The matter of dosage to the patient brings up other considerations besides those pertaining to calibration. It might be expected, for instance, that if the amount of radiation delivered to a certain point in air in a certain time were known, and if the patient were placed so that the surface of his body was at the same point, his skin would receive the same dose in the same time. Consideration of the behavior of an x-ray beam on striking a mass of material, however, shows immediately that this cannot be the case. As the beam enters the body, rays are scattered in all directions, not only once but repeatedly. Thus the paths of many of the photons are altered at all points throughout the beam. For this reason, the radiation effective at the skin surface consists of not only the direct radiation from the x-ray tube—the primary beam—but also a contribution scattered back from the underlying tissues—the back scatter.[9] Thus the skin dose is greater than the dose at the same point when the patient is not there, by the amount of this back scatter. As for any point within the body, the dose depends on the absorption of radiation in the tissues already traversed, the scattering both in this tissue and in parts beyond the point in question, and the spreading of the beam according to the inverse square law.[10] The scattering and absorption depend on the quality of the radiation and on the volume of scattering material, and the influence of the inverse square law depends on the distance from the target of the x-ray tube and the depth of the point under consideration. Hence it is obvious that here can be no simple relation between the air dose and the dose to any specified organ of the patient, but that for any particular air dose there can be a large number of skin or tumor doses. The problem is how these tissue doses, so-called, can be either measured or calculated.

The need for information about both skin and tumor doses is evident when the nature of x-ray therapy is considered. In the treatment of deep-seated lesions, the skin field through which the beam enters the body inevitably receives a larger dose than any tissue at some depth beneath the

skin. Hence, with irradiation through a single field, the tolerance of the skin is the limiting factor to the amount which can be administered. In the intensive irradiation of malignant growths, treatment can sometimes be carried out through several skin portals, the various beams being so directed as to intersect at the tumor. Even in this "cross-fire" method, it is frequently necessary to push the dose to the limit of skin tolerance for each individual field, in order to add up a large enough total dose in the deep-seated organ. Thus it is essential to know how much radiation the skin will tolerate under any special conditions of treatment, and then to know the amount it receives, so that this tolerance dose will not be exceeded.

Furthermore, as has already been stated, it is the radiation actually absorbed by the tumor cells which produces a biological reaction therein. It has just been made evident that there can be no constant relation between dose in air, or on the skin, and that reaching the tumor, hence it is evident that this latter must also be ascertained in some manner for each individual case.

It is possible to put an ionization chamber on the skin and make a measurement of the radiation reaching it. On the other hand, for points within the body an ionization chamber can seldom be introduced for purposes of direct measurement. However, it is possible to study the behavior of ionization chambers on or in masses of non-living, unit density organic material of volume comparable to that of the portion of the body within which the dose is desired. Since the body is composed of organic material of approximately unit density, it can be assumed that measurements of this type give, within reasonable limits, the same values as would be obtained in the body. This assumption is borne out by measurements made with a cadaver, and in accessible body cavities in the living patient.

Such a mass of material is called a *phantom*. Water was first used as a phantom material, but because of leakage and insulation difficulties, some other substance was highly desirable. Wax of various compositions was employed, but at present the most satisfactory material appears to be a compressed building material known as "pressdwood." This is available in sheets a few mm thick; it can be cut to any desired size and shape, and piles of such pieces readily handled. Several pieces can be fastened together, with a cavity which exactly fits around the ionization chamber, which can then be placed at any desired position within the phantom.

The ionization chambers used with such phantoms have, in general, been of air-wall material, spherical or thimble-shaped, with a volume of the order of one or two cc; they are the same chambers as are employed for determining the output of the x-ray tube in air. For surface measurements on the phantom, they are usually so placed as to be half buried; within the phantom, the depth is considered as the depth to the center of the chamber.

Measurements of this type have been made with a variety of ionization chambers, by a number of physicists. The purpose has been to find, for various combinations of voltage, filter, distance, and irradiated area, the relation of skin dose to air dose, and of deep tissue dose to either air or skin dose. The best published data vary somewhat among themselves. While the discrepancies are not beyond the accuracy desired for clinical work (in which large unknown factors are always present) their significance should be investigated.

Sources of error might be of two types, the first due to the size and shape of the chamber and the second to some sort of wall effect. When such a chamber is used on the surface of the phantom, or half-submerged, as is usually the practice, the purpose is to measure the radiation in a narrow region representing the skin surface. As is shown in Fig. 3, the geometrical relationships are rather different for the chamber and for the skin layer, and it is not certain that the ionization in the chamber actually represents that at the skin surface. Similar uncertainties exist for measurements at depths within the phantom. Hence it is desirable that some method be devised to avoid the geometrical difficulty. As far as the wall effect is concerned it causes no trouble if the chamber is "air-wall" within the range of wave-lengths present in the beam at the desired point in the phantom. This is usually the case with radiation in the usual therapy range.

For the investigation of the effect of the geometrical disturbance arising from the introduction of the ionization chamber, a special type of

FIG. 3. Diagram showing relative geometrical volumes of skin and ionization chamber.

chamber has been designed by Failla, as shown in Fig. 4.[9] This is a parallel plate chamber, but with the plates perpendicular to the axis of the beam of x-rays instead of parallel to it as in the case of the standard type. Accordingly the x-rays pass through the wall as in the case of the thimble chamber. The lower plate is continuous with the phantom; the upper is mounted on micrometer screws, so that the spacing between the two can be varied and be accurately measured. This upper (charged) electrode is completely demountable, and the chamber is provided with several interchangeable ones of different thicknesses of organic material, from a very fine silk net having an effective thickness of 0.02 mm up to one of pressdwood 3 mm thick. For all electrodes, electrical conductivity is obtained by painting with India ink or Aquadag as usual. The lower plate is constructed of Lucite or a similar substance; a fine scratch through the Aquadag separates the central area from the rest. The central part is the collecting electrode; a properly shielded wire connects it to the current measuring instrument. The outer part is grounded and serves as a guard ring. The scratch affords sufficient insulation *when a null method of measuring is employed*, so that at the instant of measuring, the collecting electrode and the guard rings are at the same potential.

With this chamber as the surface of the phantom, and a sufficiently small spacing between the electrodes, the only geometrical difference from the surface of the skin is the presence of the upper electrode. By using in turn several of different thicknesses, and constructing a curve for the readings thus obtained, it is possible to extrapolate to zero electrode thickness. In practice, for radiation in the 100–200-kv region essentially the same readings were obtained with the fine silk net and with a thin parchment. Since this indicates that there was no significant effect due to the presence of the parchment electrode, it is ordinarily employed for surface measurements.

The purpose of mounting the upper electrode on micrometers was so that the effect of even the small air space could be taken into account by taking measurements for a series of spacings and extrapolating to zero spacing. (Because of these two extrapolation possibilities the instrument was originally called an "extrapolation chamber.") However, in practice it was found that, for x-rays ordinarily employed in therapy, from the smallest possible spacing, about 0.5 mm, up to about 3 mm, the ionization *per mm spacing* was constant. Hence the extrapolation feature, which is very useful in fixing the range of spacings within which correct measurements can be made, does not have to be employed to determine the actual ionization for each position.

Within the depths of the phantom, as well as at the surface, it is important that the measuring device should not distort the geometrical relations. The extrapolation chamber is equally serviceable in this situation, if the upper electrode is a mass of pressdwood corresponding to the depth under consideration, and if varying spacings are employed so as to permit extrapolation to zero air space. In practice the upper electrode is a sheet of pressdwood, properly inked, upon which can be piled other sheets to the desired depth. The extrapolation procedure is not necessary, once the range has been determined within which the ionization per mm spacing is independent of the spacing.

For the actual determination of ionization per cc, it is necessary to know accurately the ionized volume. The diameter of the collecting electrode can be measured with a traveling microscope; the separation between the plates can be determined by measuring the electrical capacity of the chamber as a parallel plate condenser.†† It is thus possible to measure in roentgens the x-rays at any point within the phantom.

†† The micrometer screw settings cannot be used to measure the actual plate separation. They are useful in varying the spacing by accurately known amounts.

With the extrapolation chamber and an air-wall thimble chamber, parallel series of measurements of surface and depth doses have been made over a wide range of voltages (60 to 1000 kv) and of irradiated areas (5 to 400 sq. cm). They showed that, with proper precautions, measurements made with acceptable thimble chambers should not be more than a few percent in error for surface and depth dose measurements, for x-rays from 100 to 250 kv, and for fields of 50 sq. cm or more. The thimble chambers consistently give results a few percent higher than the extrapolation chamber. For smaller fields, accurate measurements become more difficult, and for very small ones impossible, with chambers of the ordinary size. (It is, of course, possible to construct special small chambers for special studies.) For higher and lower voltages, other difficulties arise, which will be mentioned below.

It has thus been shown that, within the range of voltages and fields ordinarily employed in x-ray therapy, it is practicable to calibrate the output of the x-ray tube, and to express the dose of radiation delivered to specified points within the body, in terms of a satisfactory unit based on the ionization produced in air. At the present time the vast majority of therapeutic radiologists have their machines properly calibrated. The determination of actual tissue doses is less general, partly because the necessary information has been available for a shorter time. It is becoming more widespread, and its desirability more thoroughly realized. The calculations have been facilitated by the publication of extensive tables for back scatter and depth dose for various qualities of radiation. If a radiologist has his x-ray machine calibrated for its output in air, he can determine the dose to the skin or to any particular depth in the tissue, for an exposure of a certain number of roentgens in air, by reference to these tables. Or he can measure directly the dose delivered on the skin, including the back scatter; in any case he must have recourse to tables for the dose in the deep-seated tissues.[9, 11, 12]

For x-rays generated in the neighborhood of a million volts, as has been stated earlier in this paper, it becomes necessary to increase the wall of the thimble chamber in order to be sure that the primary beam is in equilibrium with its secondaries. With the extrapolation chamber this is demonstrated by the fact that, at 1000 kv, as the thickness of the upper electrode is increased from a small fraction of a mm up to about 2 mm, the ionization in the chamber also increases as the secondary electron component of the radiation is built up to equilibrium. The significance of this situation in relation to clinical dosage is shown by the following consideration: The secondary elec-

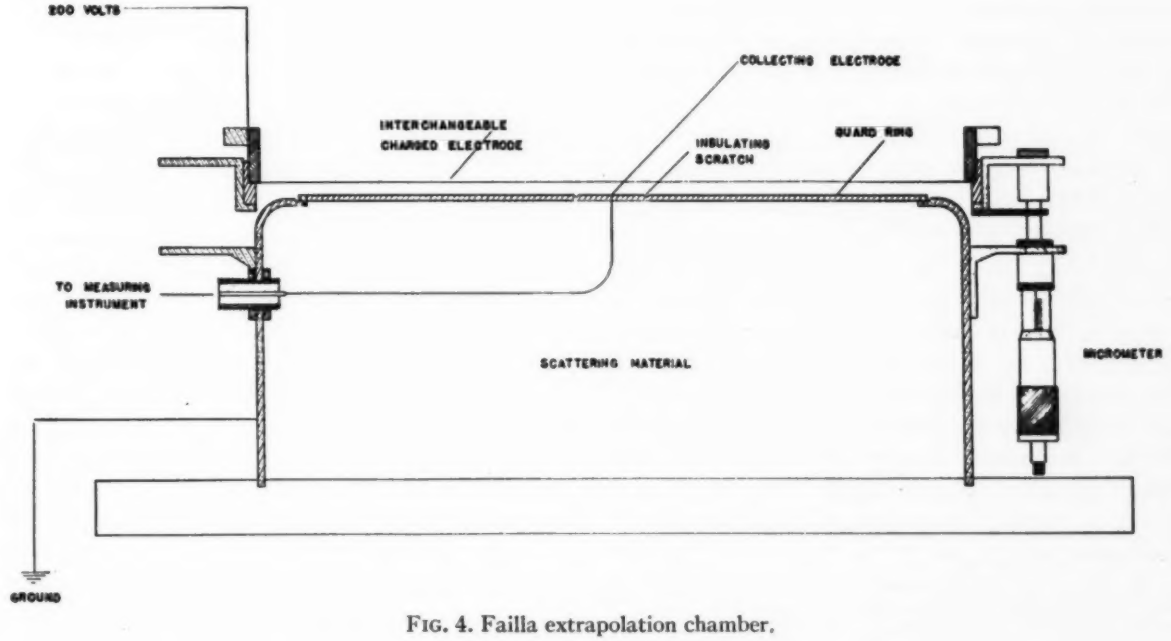



FIG. 4. Failla extrapolation chamber.

trons of 1000-kv x-rays have a considerable range in air, and, in reasonably narrow beams, such as are usually employed in therapy, many of these electrons are scattered out of the beam between the tube or the filter and the patient. The result is that at the surface of the skin the radiation does not have its full complement of secondary electrons. As the rays go through the skin, more and more electrons are produced, and since their range in tissue is very much less than in air, they remain within the limits of the beam and produce ionization there. This ionization is responsible for the biological effect, and it reaches a maximum at a depth where the x-ray beam is in equilibrium with its secondaries, which, in the case of 1000-kv radiation is about 2 mm.

It is only with the extrapolation chamber (or its equivalent), with an upper electrode whose thickness can be varied as described, that this increase in ionization with very small depths can be studied. The phenomenon would be expected to be clinically significant, for if the skin reaction observed following irradiation is due to ionization very close to the surface, there should be less reaction when equilibrium is not established until a depth greater than that of this particular region is reached. This has been verified by clinical experiments which show that, for million volt x-rays, about 20 percent less exposure is necessary to produce a skin reaction when the skin is covered with 2 mm of organic material than when it is bare. The cover provides sufficient material to bring the beam to electron equilibrium, or to provide a larger number of secondaries than exist at the surface of the uncovered skin. Accordingly in the case of these penetrating radiations, the dose delivered to the skin depends on whether or not electron equilibrium exists at the skin surface. It is evident that in specifying the surface dose with such radiation, it is essential to state what material, if any, covers the skin. Since it is desirable to keep the skin dose as small as possible, a cone to define the beam of rays, with a Bakelite closing end, such as is generally used in the 200-kv region (where this effect does not exist*) becomes undesirable in the supervoltage range. A detailed discussion of calibrations and measurements for these high voltage radiations cannot be included here.[13]

At the other extreme of x-ray quality employed in therapy, namely, at voltages in the neighborhood of 45 kv, the complication introduced by carbon photoelectrons from the wall coating of the thimble chamber has already been mentioned. The extrapolation chamber affords a method of investigating this matter. A chamber especially built for the study of these radiations has a minimum spacing of 0.2 mm. For this spacing most of the ionization in the chamber is due to the carbon electrons, whereas at spacings of a few mm most of it is due to electrons arising in the air between the electrodes. It has already been stated that because of this effect, small closed chambers are not suitable for calibration for radiation of this quality. Work on the problem of dosage measurement is still in progress; it appears that, under certain conditions the extrapolation chamber can be adapted to the determination of *relative* air and tissue doses. If then an open air chamber is used to determine the actual air dose, the tissue doses can be calculated.[14]

No consideration of dosage determination is complete without some mention of unwanted dosage—that is, the need for and determination of *protection* for everyone involved in x-ray work. These penetrating radiations are no respecters of persons; they produce biological changes in whatever organisms they reach in sufficient quantity. Hence it is essential to find out how much radiation can be tolerated over long periods of time, and how, in any particular instance, it can be determined whether this dose is being exceeded. On the basis of known exposures, a tolerance dose amounting to about 1 roentgen per week has been accepted by the International X-Ray and Radium Protection Commission.[15] This, for a 40-hour working week, amounts to $10^{-5}$ roentgen per second, or the production of about 20,000 ion pairs per second per cc of air. This degree of ionization can readily be determined by means of ionization chambers having a volume of a liter or less—there is no need for small size in this case. It may be stated in this connection that, in general, hospitals and radiologists appear to be cognizant of and attentive to the needs for protection. This is to be expected, since radiation reactions in patients are produced

* At 200 kv, electron equilibrium is established in a very small amount of matter, usually in the air intervening between the tube and the skin.

deliberately as a therapeutic measure, and are hence familiar. The seriousness of producing such reactions unnecessarily is appreciated. In educational institutions and manufacturing plants, where no patients are treated and radiation reactions are usually unknown, the dangers are not so apparent. With the increased employment of powerful x-ray machines and of neutron generators and artificially radioactive substances in industry and in university laboratories, the need for adequate protection cannot be too frequently emphasized.


**BIBLIOGRAPHY**

(1) G. Failla, Radiology **13**, 293 (1929).
(2) G. Failla, Radiology **29**, 202 (1937).
(3) Recommendations of the International Committee for Radiological Units (Chicago, 1937); Radiology **29**, 634 (1937).
(4) G. Failla, Am. J. Roent. and Rad. Therapy **21**, 47 (1929).
(5) Hermann Behnken, Strahlentherapie **26**, 79 (1927).
(6) Lauriston S. Taylor, Radiology **15**, 49 (1930).
(7) Hugo Fricke and Otto Glasser, Am. J. Roent. and Rad. Therapy **13**, 453 (1925).
(8) Walter Friedrich, Am. J. Roent. and Rad. Therapy **40**, 69 (1938).
(9) Edith H. Quimby, L. D. Marinelli, and J. H. Farrow, Am. J. Roent. and Rad. Therapy **39**, 799 (1938).
(10) Edith H. Quimby, Murray M. Copeland, and Robert C. Woods, Am. J. Roent. and Rad. Therapy **32**, 534 (1934).
(11) Edith H. Quimby, Radiology **38**, 261 (1942).
(12) Edith H. Quimby and George C. Laurence, Radiology **35**, 138 (1940).
(13) Edith H. Quimby and Elizabeth F. Focht, Am. J. Roent. and Rad. Therapy **46**, 376 (1941).
(14) Edith H. Quimby and Elizabeth F. Focht, Am. J. Roent. and Rad. Therapy, to be published in 1943.
(15) International Recommendations for X-Ray and Radium Protection; Radiology **30**, 511 (1938).


# New Books

## Introduction to Modern Physics

By F. K. Richtmyer and E. H. Kennard. Third edition. Pp. 723+xv, Figs. 234, 16×23½ cm. McGraw-Hill Book Company, Inc., New York and London, 1942. Price $5.00.

E. H. Kennard undertook the worthy task of re-editing Richtmyer's well-known *Introduction to Modern Physics*. Since the appearance of the last edition (1934) the experimental advance particularly in the physics of the nucleus and of elementary particles, as observed in the cosmic radiation, as well as the expansion of applications of wave mechanics, has been so rapid that extensive additions and changes were necessary if the book was to remain true to its title.

"Briefly, the course of the revision has been as follows. The historical introduction, abbreviated, is followed by a single chapter, rewritten, on those topics in electromagnetism which are needed for the subsequent discussion and are not always adequately treated in general textbooks. Between the chapters on the photoelectric effect and on the origin of the quantum theory is inserted a short chapter on relativity. Then comes a single chapter, replacing Chapters IX and X, containing the essential ideas concerning the nuclear atom, spectral series, and atomic quantum states. The Bohr theory of hydrogen is retained, because of its pictorial value, but with a clear statement as to its true status. A descriptive chapter on wave mechanics is then followed by a single chapter, replacing Chapters XI and XII, on the theory of the periodic table and on optical spectroscopy. The chapter on specific heats, which might perhaps have been omitted, is inserted next. The chapter on x-rays has been thoroughly revised in collaboration with Professor L. G. Parratt. The book closes with the chapter on the nucleus, considerably extended, and a new chapter on cosmic-ray phenomena."

It can certainly be said that the goal of producing an up-to-date equivalent of the first edition has admirably been reached. One might say that the emphasis is again laid on the physics of particles and rays, so that the student is left with the impression that modern physics has about completely deserted the field of physics of matter in bulk and can do without such old-fashioned concepts as entropy, Nernst's theorem, kinetic theory of matter, etc. From that standpoint the addition of a short chapter on thermodynamics and statistics might have appeared more urgent than that given on relativity. However, it is certainly up to the author to make his choice and it is mostly an advantage if a personal standpoint finds its due expression. The new chapters on cosmic radiation and nuclear physics give very valuable elementary representations of the status of the respective subjects up to 1940 and maintain the standard established by the former edition.


F. London
*Duke University*


★

Attention is called to Special Paper No. 36 "*Handbook of Physical Constants*" published January 31, 1942 by The Geological Society of America. The 325-page volume, edited by Francis Birch, contains 21 sections prepared by 19 specialists cooperating through the Divisions of Chemistry, Physics, and Geology of the National Research Council. Copies may be consulted in geological libraries and are available in paper covers at a price of $1.40 on application to The Geological Society of America, 419 West 117 Street, New York, New York.

# Frictional Phenomena. XI

## C–Solids

BY ANDREW GEMANT

*The Detroit Edison Company, Detroit, Michigan*

### Chapter XI. The Plastic Flow of Solids

***Abstract***

THIS chapter deals with (a) the general features of plastic phenomena in solids, (b) the so-called Prandtl model, giving a basic and pictorial explanation of plastic processes, (c) the physical theories of R. Becker and G. I. Taylor, bearing a closer relation to physically measurable quantities, and (d) some experimental results illustrating the chief features of plastic flow in solids.

### *41. General Relations Observed in Plastic Flow*

The flow properties of solids under steady loads are called plastic properties, and the present chapter will be devoted to this particular field. It is the logical extension of the viscous phenomena of gases and liquids, which were dealt with in the first two parts of this monograph. Actually the field of plasticity is so wide that whole books have been written on it. It will be understood, therefore, that the present discussion is not intended to be exhaustive. The chief aim will be the presentation of the fundamental experimental results and theories, particular attention being paid to a continuity of the discussion with regard to the preceding parts.

The multitude of processes occurring when solids are subjected to stresses can best be illustrated by means of the schematic diagram in Fig. 80. The stress can be either tension or shear, according to the experimental conditions. However, since shear lends itself easily to physical and mathematical considerations, it will be best to consider a simple shearing motion in connection with the diagram. It is assumed that a given stress is suddenly applied to a specimen and that the deformation as a function of time is measured. Two types of flow are generally observed; that type represented by the lower curve will be considered first.

First, a sudden deformation occurs, its extent being determined by the elasticity of the material. This part ($AB$ on the diagram) is completely reversible, disappearing instantaneously upon removal of the load. A second deformation then takes place ($BC$) that is characterized by a steadily decreasing rate. The decrease of the rate is roughly exponential, but measurements reveal a pronounced deviation from the exponential. This part, sometimes called first stage of creep, is also reversible upon removal of the load. This portion of the deformation obviously is not due to the elasticity, as described by Hooke's law, since elastic deformations are instantaneous. It is the initial rapid stage of the plastic deformation of solids—that slow process that is called plastic flow or creep.

The next part of the plastic flow ($CD$) is characterized by a constant rate. This part obviously bears the greatest similarity to viscous flow in gases and liquids, since they, too, are distinguished by a constant rate. If the stress is divided by the gradient of the portion $CD$, a quantity is obtained that is a direct analogue to the viscosity $\eta$, as defined in the previous sections. The chief difference is that this quantity is not a constant of the material, as is the viscosity in the case of gases and liquids (see definition of viscosity constant in Section 3). In other words, the stress *vs.* rate of shear relation is not linear, i.e., characterized by a single viscosity constant. Cases of non-linearity were encountered in gel-like liquid structures (Chapter VIII). In solids, non-linearity between stress and rate of shear is not an exception, as in the case of liquids, but the rule. Purely qualitatively, the relation in this case is similar to that in non-Newtonian liquids, namely, such that the rate of shear increases more rapidly than does the stress. Or, expressed in another manner, the viscosity decreases with increasing stress (and rate of shear). The function according to which this increase takes place will be discussed below in greater detail.

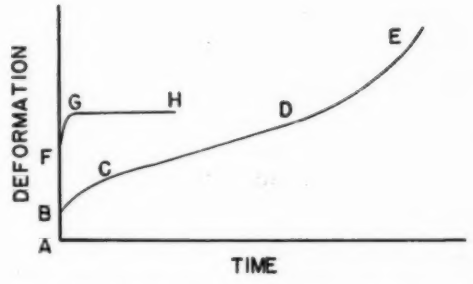



FIG. 80. Diagram of plastic deformation *vs.* time (lower curve: creep, upper curve: slip).

Because of the fact that there is no constant viscosity in solids (compare with Section 3, Chapter I), other names, such as flow resistance or creep resistance, are often used in the literature. The phase *CD* is sometimes referred to as the second stage of creep.

A third stage that often follows after a long time, depending upon the stress, is indicated by the line *DE*, in which the flow becomes more rapid, ultimately leading to fracture. This latter stage will not be discussed further in this monograph.

The plastic flow as exhibited by the whole curve is often referred to as creep, because of the slowness of the major constant rate portion. Non-crystalline solids, such as glasses and organic plastics, exhibit only this particular type of plastic flow.

Another type of plastic flow is indicated schematically by the upper curve in Fig. 80. Here *AF* indicates the elastic displacement which is then followed by a fairly rapid flow *FG* that very soon stops, the final deformation *H* being essentially independent of time. In releasing the stress, only the elastic part *AF* is reversible; part *FG* remains as a more or less permanent deformation. The chief difference between this type of flow and the previous one is that, owing to secondary complications, the flow resistance rapidly increases so much so that the flow automatically stops, at least practically. The material is now obviously harder than it was originally, an effect called strain hardening. This process is called slip, in contradistinction to creep, because of the relative speed with which it takes place, and the abruptness with which it stops.

With crystalline materials, particularly single crystals of metals and inorganic solids, the usually observable flow is the slip. At sufficiently low stresses and with correspondingly more sensitive apparatus the branch *FG* (deformation progressing with time) is sometimes detected. While this branch itself is sometimes referred to as creep, the whole process, if it comes to a standstill, is essentially of the slip type.

Polycrystalline metals take an intermediate position between single crystals and glassy materials. Accordingly their plastic behavior is somewhere between those of the two extreme groups. While the usually observable flow of polycrystalline metals is the creep, pronounced effects of strain or work hardening do exist.

In the presentation shown by Fig. 80 the stress was kept constant, and the time was varied. Another frequently used presentation is that by which a given time is chosen and the rate of deformation is plotted against stress. As rate of deformation one may choose the differential gradient of the branch *CD*, or, as is frequently done in technical procedure, the total deformation attained during, say, the first 1000 hours. This second procedure will yield somewhat higher rates than the first physically correct method, but for technical purposes the difference is of no great importance.

Such a diagram is shown in Fig. 81. As can be seen, up to a certain point *A* the rate is practically zero. This value of the stress is often called yield value, as it is called in the case of structurally viscous liquids. Theoretically this point

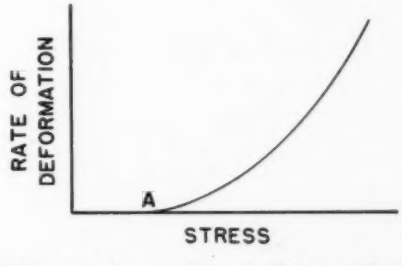



FIG. 81. Diagram of rate of plastic deformation *vs.* stress.

has limited significance only, since there is little doubt that the rate for smaller stresses is not zero but merely is too small to be observed with the particular apparatus in use. Thus, the yield value is actually a function of the sensitivity of the measuring device and therefore it is more of technical than of physical importance. The in-

crease of the rate with increasing stress is often nearly exponential, as will be discussed below.

The yield value of metals—steel, for instance—is often considered in designing structures, it being assumed that a structure becomes unsafe if the stress at any point reaches the yield value. Incidentally, Baker and Roderick[1] recently showed that rigid structures can safely carry stresses beyond the yield value. Considerations of this kind are not affected by the definition of the yield value—whether it exists in an absolute sense or only relatively, as defined above.

It is interesting to note the similarity between the curves Figs. 80 and 81 with corresponding curves on dielectric materials if the mechanical stress is replaced by the electric field intensity, and the rate of deformation by the current density. In Fig. 80, *AB* represents the instantaneous charging current of the condenser, *BC* the reversible absorption current, *CD* the continuous d.c. leakage current, *DE* a branch leading to electric breakdown, and *GH* a condition in which the ion content is depleted. In Fig. 81 the branch up to point *A* corresponds to very minute currents at small field intensities, and the rest of the curve is the usual exponential increase of the current which is due to several possible causes. This analogy is quoted here because of its instructiveness, without a connection being assumed between the two processes. However, certain connections undoubtedly exist (see, for instance, the explanation of ionic mobility in liquids by means of the viscosity concept, Section 38, the discussion of ionic mobility in solids to appear in Section 48, and the discussion of dielectric losses in solids to appear in Chapter XIV).

### *42. The Prandtl Model of Plastic Properties*

In this section a theoretical explanation of the plastic properties just outlined will be given by means of the so-called Prandtl[2] model. In the next section some other physical theories, aimed at the same purpose, will be sketched briefly. They are by no means in contradiction to the Prandtl model. The latter is a theory more pictorial than physical, using concepts on the structure of solids that are basically factual even if not strictly related to experimental structural data. Thus it is a model particularly suited for the understanding of plastic properties by those that do not care to go too deeply into detail. That is the reason for which it is incorporated here.

We consider (see Fig. 82) a plane, indicated by the line *AB*, across the plastic material, this plane being the seat of a shearing motion as shown by the double arrow at the right. It is assumed that the lower portion moves towards the right with a velocity or rate of shear *u* relative to the upper portion. The problem is to find the stress pertaining to this rate of shear, the ratio of the two giving the viscosity, or flow resistance.

Molecules of the lower portion will be considered, and the forces acting upon them will be computed. In first approximation these molecules can be assumed to be equidistant as shown by row *1*. The forces acting upon this row will originate partly from the molecules of the lower, partly from those of the upper portion. Since the

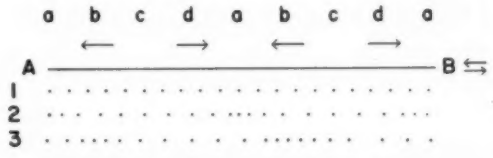



FIG. 82. Prandtl's model of plastic flow.

lower portion shifts as a whole against the upper one, it is safe to assume that the retarding forces that must be overcome by the external stress will be those from the upper portion. It is those forces with which we are particularly concerned in this problem. Besides, only one component of those forces will be taken into account, namely the component parallel to the plane of shear, since the external stress is in the same direction.

The forces originating from the upper portion are idealized as having periodic attraction sources located at points like *a*. In the points between, as indicated by *c*, we then can assume sources of repulsion. The physical nature of these sources is left open in this model; they might be molecules, or crystallites of any size. The component of these forces parallel to the plane of shear is then represented by the arrows located above the plane at *b* and *d*. Prandtl, of course, formulated these components mathematically, but these details are not discussed here.

As a result of the forces outlined above, the row of molecules when at rest will assume an equilibrium position as shown by row *2*, charac-

terized by periodic changes of density. At locations $a$ the density will be above the average, at locations $c$, below the average.

In considering a shift to the right of row *2* against the upper portion of the material, one must realize that the structure of the row (as of all molecules of the material) is mobile rather than rigid. The molecules vibrate in accordance with their kinetic thermal energy, and the equilibrium positions as shown in row *2* are, in reality, only the centers of the vibratory motion. If, therefore, the row moves to the right, an immediate readjustment of the row might be expected, to the end that the individual molecules below $a$ will move apart as they pass through $c$, then again move closer to each other as they pass the next point $a$ and so on. In other words, one might expect that the distribution of density in row *2* with regard to the arrows above will be the same in motion as in rest.

This statement, however, has to be modified. It would be correct, if the heat motion of the molecules had such an intensity that immediate readjustment could be assumed, in other words, at very high temperatures. For normal temperatures this process of readjustment will take a little time, and accordingly the distribution of density in the row with regard to the arrows will be shifted a little towards the right, as shown by row *3*. The high densities in that row are shown to be below points $b$, and the low densities below points $d$. A definite shift between the rest distribution *2* and the moving distribution *3* has taken place, that will be denoted by $\xi$ and which in the specific example equals the distance $ab$.

This shift $\xi$ will evidently be a function of the rate of shear $u$, as can be seen by the following reasoning. The shift $\xi$ represents a mobile equilibrium, a stationary condition, characterized by a small increase in the unit time, due to the shearing motion, and a small decrease in the unit time, due to the heat motion of the molecules, such that these two figures, the small increase and the small decrease, are equal. This condition yields an equation that will show the shift $\xi$ as a function of $u$. The increase just mentioned is obviously proportional to $u$, and the decrease is proportional to the product of two factors: an "intensity factor" $f(T)$, where $T$ = absolute temperature, and a "capacity factor" $\varphi(\xi)$. Both $f$ and $\varphi$ must increase with the argument of the function, although the nature of this increase can be deduced only mathematically. The general relation will be:

$$u = f(T)\varphi(\xi). \tag{119}$$

The force that pulls row *3* backwards will be a function of $\xi$, too, as can be seen from the figure. For $\xi = 0$ the force is zero, since it corresponds to a stable equilibrium. With increasing $\xi$ the force to the left will increase. This force is equal to the external stress $S$ that causes the material to flow with rate of shear $u$. Since, then, $S$ is a function of $\xi$, Eq. (119) may be written

$$u = f(T)\psi(S), \tag{120}$$

where $\psi$ is a function increasing with $S$.

In carrying out the computations involved three different cases can be distinguished.

(a) Small shear. For values of $u$ and $S$ below certain limits (to be specified under $c$) it is found that $\psi$ is a linear function of $S$; in other words, the viscosity will have a constant value. The function $f$ is an exponential function of the absolute temperature. In solving the equation for $S/u$, a quantity proportional to the viscosity, the result is:

$$S/u = (\pi G/8\nu)e^{B/RT}. \tag{121}$$

Here $G$ = shear modulus of elasticity, $\nu$ = frequency of thermal vibrations, $B$ = molar energy (average value) required for the molecules to jump back into the equilibrium position (say, from $b$ to $a$). The existence of an activation energy $B$ follows from a detailed picture of the potential levels which shows that an energy hump has to be overcome in order that a molecule be able to jump back into the equilibrium position.

It can be seen that high temperatures cause low viscosities, since the attained displacement (between rows *2* and *3* in Fig. 82) will be small; hence the force pulling back is small, too. Plastic resistance to flow decreases with temperature. Equation (121) is the analogue to Eq. (66) in Section 20, valid for liquids. There the compressibility $\kappa$ replaced the shear modulus $G$.

(b) High shear. The model discussed here indicates that there must be a maximum value of displacement $\xi$ between the rest and the stationary distribution (rows *2* and *3*), namely, the distance $ac$. This also means that there must be a

maximum value for the stress, since the latter increases with the displacement $\xi$. The theory indicates that this maximum stress in plastic flow must be of the order $G$, the shear modulus.

(c) Intermediate shear. Most interesting are the conclusions to be drawn regarding the intermediate range between (a) the linear range and (b) the limiting stress. As to the rate of shear *vs.* stress relation, the function $\psi(S)$ in Eq. (120) turns out to be approximately an exponential function of $S$. The physical reason of this exponential relation is of the same nature as that discussed below in connection with Becker's theory (Eq. 124). Solving for $S$ in Eq. (120), an equation of the form

$$S \sim \log u + \text{const} \qquad (122)$$

is obtained which means that the stress increases only logarithmically with the rate of shear. The flow resistance in this range decreases with increasing rate of shear, approaching asymptotically the value zero for the limiting case. This type of relation is confirmed by experiment, as will be seen later. The model thus gives for the $u(S)$ relation a curve as shown in Fig. 83, with $AB$ the linear and $BC$ the logarithmic range. The abscissa corresponding to $C$ is the limiting stress. No actual yield value appears in this curve. Its existence in the general diagram, Fig. 81, is, as previously mentioned, probably due to a threshold in the sensitivity of the measuring device.

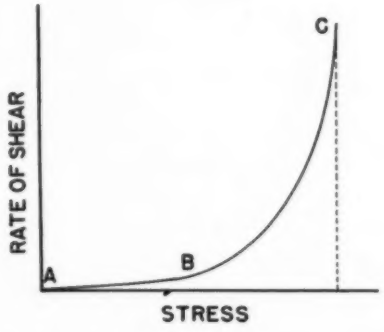



FIG. 83. Diagram of rate of plastic shear *vs.* stress according to Prandtl's model.

The model also indicates the values of the rate of shear at which the linear portion $AB$ gives way to the logarithmic portion, i.e., at which the viscosity will cease to be a constant. This transition is controlled by the relation:

$$u/\delta\nu \leqq RT/Be^{B/RT} \qquad (123)$$

($\delta$ = a distance of the order $ab$ in Fig. 82). For rates satisfying the upper sign the $u(S)$ relation is linear; when the rate becomes large enough for the lower sign to hold, the $u(S)$ relation will become logarithmic. The right side of (123) is proportional to the ratio of the thermal to the potential energy of the molecules. Thus, for high ratios the limiting rate is high and the plastic material (pitch, for instance) behaves similarly to a liquid up to high rates of shear; materials such as metals, having a low ratio, will not have a constant viscosity even for low rates of shear. It is instructive to compare Eq. (123) with the corresponding equation (Eq. 95) of Eyring's theory which bears, as already mentioned, a certain similarity to Prandtl's model.

Apart from the steady state stage ($CD$ in Fig. 80) the model also demonstrates the initial stages ($AB$ and $BC$). By means of the model it is possible to set up the equations that are valid for the sudden unloading of a system in a stationary flow. It then appears that two processes take place. One is an instantaneous recovery corresponding to a shear of $S/G$, as required by the elasticity of the material. The other, a type of relaxation process, is due to the displacement $\xi$ as shown in Fig. 82 and this recovery disappears at a finite rate, corresponding to the thermal motion of the molecules. The displacement $\xi$, maintained by the flow, gives rise to an internal elastic stress, when the external stress is relieved, and this stress causes the displacement to disappear with a speed prescribed by a flow resistance similar in nature to but much smaller than the stationary flow resistance, discussed above. An ideal so-called Maxwellian body possessing elasticity and viscosity (discussed in reference to gels in Section 34) exhibits exponential relaxation; the Prandtl model, on the other hand, indicates a relaxation that, compared with the exponential, slows down with time. This also is in keeping with observations.

### *43. Physical Theories of Plastic Flow*

While the Prandtl model can be considered as a basis for further theories, other authors have tried to establish theories based on assumptions

having a more evident physical background. In particular A. Smekal[3] stressed the importance of flaws or loose spots present not only in polycrystalline materials but also in single crystals. While plastic flow in a perfectly ideal lattice would be possible only at very high stresses (a conclusion theoretically arrived at and contrary to experience with real crystals) the presence of such flaws explains how flow can be initiated at much lower than the theoretical stress. It is possible to correlate the occurrence of flaws to the periodic variation of the force considered in the Prandtl model.

R. Becker[4] and later E. Orowan[5] developed a simple and attractive theory along these lines. Plastic flow is considered as a sum of individual jumps, occurring at faulty spots. The physical reality of this assumption is strengthened by observations of Joffe who found that the shearing of rocksalt is accompanied by a large number of audible ticks, indicative of the stepwise occurrence of the shear.[6]

It is now assumed that there is a theoretical stress $S_{th}$ at which shear is possible, this stress being computable on the basis of the lattice forces. Whenever the kinetic energy locally exceeds the theoretically required energy, a local jump will take place. The probability of such jumps is particularly large at flaws where, for geometric reasons, a concentration of the applied stress $S$ takes place with the effect that the local stress will be $qS$ with $q>1$. Rate of shear is now determined by the frequency of jumps at flaws, a frequency that can be computed by means of the Boltzmann theorem, by substituting the following term for the necessary potential energy to be accumulated:

$$v(S_{th}-qS)^2/2G,$$

where $v$ is the volume of the flaw. One then has for the rate of shear

$$u=C\exp\left[-v(S_{th}-qS)^2/2GkT\right] \qquad (124)$$

with $C$=a constant. It can be seen that both increasing stress and increasing temperature cause the rate of shear to increase. Solved for $S$, we obtain

$$S=\frac{1}{q}\left[S_{th}-\left(\frac{2GkT}{v}\log\frac{C}{u}\right)^{\frac{1}{2}}\right]. \qquad (125)$$

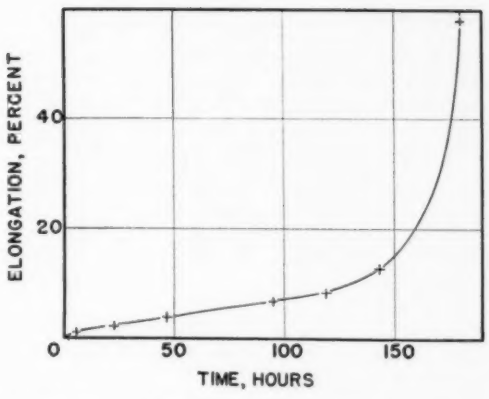



FIG. 84. Elongation-time curve for the plastic flow of lead; nominal tensile stress 35 kg/cm², temperature 66°C (according to Moore, Betty, and Dollins).

This equation shows a logarithmic variation of the stress with the rate of shear as in Eq. (122), and as shown by experience.

$S_{th}$ for metals is of the order $10^{10}$ dyne/cm². From experimental results on the flow of tungsten, copper, and cadmium values for $q$ ranging from 1 to 1000 could be deduced. From the computed value for $v$ the linear dimension of a flaw appears to be $4\times10^{-7}$ cm for tungsten.

If a yield value is explained on the basis of a threshold value for observable rates of shear, then $u$ in Eq. (125) can be replaced by this particular threshold value, in which case the equation can be interpreted as a formula for the yield value of the stress $S_0$. One has

$$S_0=(S_{th}/q)-K\sqrt{T} \qquad (126)$$

with $K$=a constant, the value of which obviously depends on the sensitivity of the measuring apparatus. Experience verifies that the yield value decreases with increasing temperature in a parabolic manner. No attempt has been made to explain hardening phenomena on the basis of this theory.

Recently deBruyne[7] deduced an equation between rate of shear and stress, claimed to be valid for all the main types of plastic and viscous flow.

Another theory, developed by M. Polanyi and G. I. Taylor[8] explains particularly the slip type of plastic flow and shear hardening, but is not so much concerned with flow resistance. Plastic flow is again explained as the sum of individual displacements occurring at flaws. A flaw is specified as a spot in which the density of the molecules

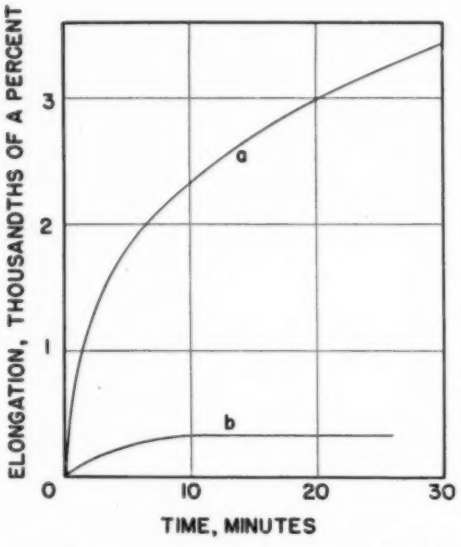



FIG. 85. Elongation-time curve for the plastic flow of polycrystalline tin; (a) extension, (b) recovery. Shearing stress 30 kg/cm² (according to Chalmers).

(at rest!) in half of its volume is slightly above normal and in the other half slightly below normal, average density being maintained. Owing to this density distortion each flaw is the seat of internal stresses. In particular two separate flaws of opposite orientation, called a pair of dislocations, will try to merge or, so to speak, to cancel each other: there is a force of attraction active between them. External shearing stresses, on the other hand, will under certain circumstances pull them apart in preferential, crystallographically determined directions, leaving in the end the interior of the crystal at its normal density, but causing a slight accumulation of molecules on two opposite surfaces of the crystal, a process equivalent to a plastic displacement of about one molecular distance. The theory of elasticity shows that the external shear will overcome the natural tendency of flaws to merge only if the shear is greater than $G\lambda/2\pi h$ where $\lambda$ = molecular distance and $h$ = separation of the two flaws in a direction perpendicular to the shear. This value can be interpreted as the yield value of the pair of dislocations. The Taylor theory, in contrast with the Prandtl model, or the theory of Becker leads to a real yield value, independent of the sensitivity of the measuring apparatus.

In the further development of the theory, the problem is considered as being two-dimensional, assuming dislocation lines all parallel to each other. Only a two-dimensional cross section of the crystal (normal to the dislocation lines) must then be considered. The number of pairs per unit surface $n$ is assumed to increase with increasing shear. It is as if the crystal would deviate from the ideal more and more with increasing shear. This number, or density of dislocations $n$ influences both the yield value and the shear that will take place when the yield value of stress is exceeded.

The distance $h$ in the above mentioned expression will be obviously proportional to $1/\sqrt{n}$; thus the yield value $S_0$ becomes

$$S_0 = CG\lambda\sqrt{n}. \tag{127}$$

The displacement resulting from the separation of one pair of dislocations is $\lambda$, in case the pair is pulled apart completely; but only $\lambda l$ (for a cross section of 1 cm²) if the path of travel of the pair is $l$ and not a full cm. The distance $l$ is a kind of mean free path of the dislocations. The product $\lambda l$ is also the shear that takes place owing to the stress $S_0$. Since there are $n$ pairs in 1 cm², the total shear $\zeta$ will be

$$\zeta = \lambda l n. \tag{128}$$

The influence of $n$ on both $S_0$ and $\zeta$ is thus shown by Eqs. (127) and (128). Eliminating $n$, we get

$$S_0 = CG(\lambda\zeta/l)^{\frac{1}{2}}, \tag{129}$$

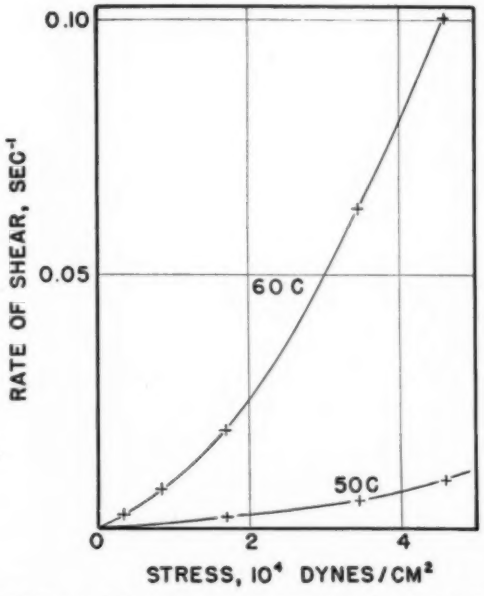



FIG. 86. Rate of shear *vs.* stress curves for the plastic flow of Chatterton compound (after Braunbek).

indicating a parabolic relation between the shear of a crystal and the yield value. This is the experimentally found relation controlling shear hardening, as will be shown in Section 44.

The order of magnitude for $l$ as derived from Eq. (129), using experimental data, is $10^{-4}$ cm for copper and iron at room temperature, $10^{-3}$ for NaCl at around 100° C, and $10^{-2}$ cm for NaCl at around 500° C.

If the assumption is made that the distance $l$ is not the result of one instantaneous jump, but of a slow movement having the velocity $v$, then Eq. (128) leads to an equation for the rate of shear $u$:

$$u = \lambda v n. \tag{130}$$

Recent calculations on the dislocation theory, claimed to be in numerical agreement with experiments, were carried out by Koehler.[9]

In conclusion it should be emphasized that the mechanisms of viscous flow in liquids and plastic flow in solids are similar, as is indicated by the similarity of the equations. The essence of the mechanism is this: an external stress brings forth a component in its own direction of the normally irregular molecular motion, consisting in jumps over potential barriers due to the mutual attraction of molecules. There are two chief differences, however, namely:

(1) A liquid flows at any point, whereas the flow in solids appears to be located near flaws.

(2) A liquid flows in any direction, whereas flow in solids, particularly crystals, is restricted chiefly to certain crystallographically determined directions.

For further information on this subject the reader is referred to a recent extensive review by Seitz and Read.[10]

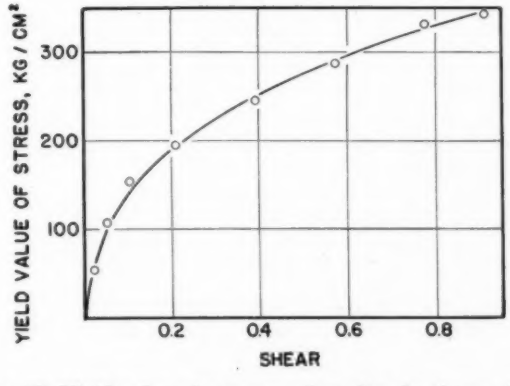



FIG. 87. Yield value of stress *vs.* shear for aluminum single crystal at room temperature (according to G. I. Taylor).

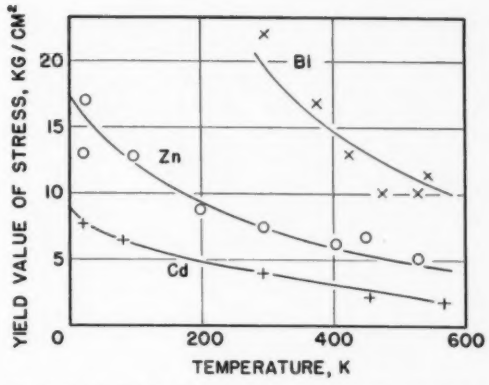



FIG. 88. Yield value of stress *vs.* temperature for single crystals of Bi, Zn, and Cd (according to Schmid).

### 44. *Experimental Material Illustrating Plastic Flow*

In this section the chief facts related to plasticity, as discussed in the three preceding sections, will be illustrated by means of a few curves, representing experimental results.

First some curves showing the first type of plastic flow, called creep, will be presented. The elastic displacement, shown in Fig. 80, is usually not included in experimental curves.

Figure 84 is a typical creep curve of commercial lead used for cable sheathing, as obtained by Moore and co-workers.[11] The three stages—decreasing, constant, and increasing rate—are clearly visible. More will be said of this particular research in the next chapter in the discussion of the plastic properties of lead as cable sheathing.

Figure 85 illustrates the plastic flow (creep type) of polycrystalline tin, after measurements of Chalmers.[12] Only stages 1 and 2 are noticeable. Curve $a$ shows the deformation upon loading, curve $b$ shows the recovery upon removal of the load. Curve $b$ is not identical in magnitude with the variable rate component of curve $a$, as should be the case if this part were fully reversible as suggested by Prandtl's model. In other instances the reversibility is much more pronounced than here.

From the straight portion of elongation-time curves the rate of deformation is determined; this can then be plotted as a function of some variable like stress or temperature. Such curves are shown in Fig. 86, referring to Chatterton compound (a

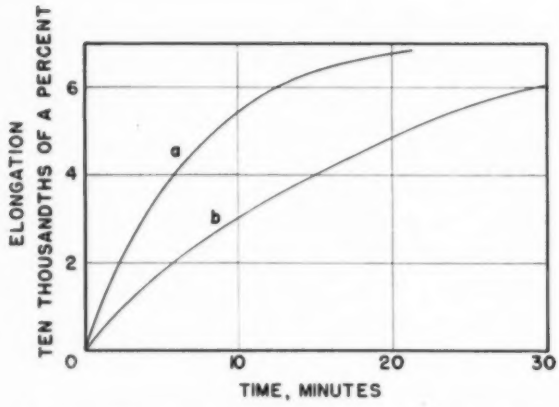



FIG. 89. Elongation-time curves for the plastic flow of *a* Williams-Harvey tin, *b* Chempur tin (both single crystals). Shearing stress 12 kg/cm² (after Chalmers).

mixture of bitumen and gutta percha), as obtained by Braunbek.[13] It can be seen that only in the first portion of the curves does the rate increase proportionately to the stress, as indicated by a constant flow resistance [Eq. (121)] in Prandtl's model. For higher stresses the resistance decreases, as predicted by Prandtl's model [Eq. (122)] and by Becker's theory [Eq. (125)].

Next some curves are given that illustrate the behavior of slip in single crystals of metals.

Figure 87 shows how strain causes hardening, that is, an increase of the yield value, necessary to cause further deformation. It refers to an aluminum single crystal in shear, obtained by G. I. Taylor. It was shown in Section 43 how Taylor's theory accounts for the parabolic relation between yield value and shear according to Eq. (129).

Figure 88 presents yield values for single crystals of cadmium, zinc, and bismuth as a function of temperature, according to Schmid.[14] The yield value decreases with increasing temperature in a parabolic manner. This parabolic relation is explained by Eq. (126), derived from Becker's theory.

Figure 89, after Chalmers, shows the kind of flow observable on single crystals at relatively low stresses. The chief difference between this flow and that of the creep type observable with polycrystalline materials and plastics is that the flow resistance increases continuously until the flow stops. This can be seen clearly on curve *a* of Fig. 86.

These curves are presented here as being representative of the chief plastic phenomena. Further experimental material will be given in the next chapter in connection with technical applications of plasticity.

Measuring equipment, so-called plastometers, will not be discussed here systematically. A good review will be found in an article by Nieuwenburg.[15] Bailey[16] discusses a new instrument utilizing the torsion principle, and claimed to be very versatile. Kinney[17] describes how the so-called Gardner mobilometer can be calibrated in absolute terms by means of liquids of known viscosity. Debing and Silberkraus[18] discuss the different methods for measuring plastic resistance of phenolic molding powders, and conclude that no single method is completely adequate for practical requirements. Several instruments will be described in the next chapter.

## *Bibliography for Chapter XI*


1. J. F. Baker and J. W. Roderick, Trans. N. E. Coast Inst. Eng. Shipb. **57**, 1 (1940–41).
2. L. Prandtl, Zeits. f. angew. Math. and Mech. **8**, 85 (1928).
3. A. Smekal, Physik. Zeits. **34**, 633 (1933).
4. R. Becker, Zeits. f. tech. Physik **7**, 547 (1926).
5. E. Orowan, Zeits. f. Physik **98**, 382 (1936).
6. W. G. Burgers and J. M. Burgers in *First Report on Viscosity and Plasticity, second edition* (Nordemann, New York, 1939), p. 192.
7. N. A. deBruyne, Proc. Phys. Soc. **53**, 251 (1941).
8. G. I. Taylor, Proc. Roy. Soc. **A145**, 362 (1934).
9. J. S. Koehler, Phys. Rev. **60**, 397 (1941).
10. F. Seitz and T. A. Read, J. App. Phys. **12**, 100 (1941).
11. H. F. Moore, B. B. Betty, and C. W. Dollins, Univ. Illinois Bull. **35**, No. 102 (1938).
12. B. Chalmers, J. Inst. Metals **61**, 103 (1937).
13. W. Braunbek, Zeits. f. Physik **57**, 501 (1929).
14. E. Schmid, Proc. 3rd Int. Congress App. Mech. **2**, 249 (1930).
15. C. J. v. Nieuwenburg in *Second Report on Viscosity and Plasticity* (Nordemann, New York, 1938), p. 241.
16. C. R. Bailey, Ind. Eng. Chem., Anal. Ed. **13**, 173 (1941).
17. P. W. Kinney, Ind. Eng. Chem., Anal. Ed. **13**, 178 (1941).
18. L. M. Debing and S. H. Silberkraus, Ind. Eng. Chem. **33**, 972 (1941).

# Here and There

### Recent Staff Changes

Professor Ariel A. Benedict has resigned from the Department of Physics at Iowa State College to become Head of the Department of Physics at Muskingum College, Ohio.

Howard W. Eves, formerly on the mathematics staff of the Tennessee Valley Authority at Chattanooga, Tennessee, has been appointed Assistant Professor of Applied Mathematics at the College of Applied Science of Syracuse University. Kenneth C. Tippy has been appointed Assistant Professor of Civil Engineering at the same college.

Henry Eyring, Professor of Chemistry at Princeton University, will join the staff of the Department of Chemistry at the Polytechnic Institute of Brooklyn during the present academic year as Visiting Professor of Physical Chemistry.

Professor Howard M. Fry has been appointed Head of the Department of Physics at Franklin and Marshall College.

Professor Archie Higdon of the Department of Theoretical and Applied Mechanics at Iowa State College has been called to service as a major with the Army Air Forces at Jefferson Barracks, Missouri.

Dr. Charles Byron Jolliffe, Assistant to the President of Radio Corporation of America and Chief Engineer of RCA Laboratories, has been appointed Vice President and Chief Engineer of the RCA Manufacturing Company, Camden, New Jersey.

Charles A. Mabey, physicist at the Bristol Company, Waterbury, Connecticut, has recently been appointed Director of Research of that firm.

It is reported that Professor W. Peddie has retired from the Harris Chair of Physics at University College, in the University of St. Andrews, Dundee, Scotland.

James W. Schade, Research Director of the B. F. Goodrich Company until December, 1941, has joined the staff of the University of Akron to give a special course in rubber technology in cooperation with the Firestone, Goodrich, and Goodyear Rubber Companies. The lectures will be especially designed for young men entering industry, aiming to give them a broad understanding of the entire rubber production field.

### Acheson Medal Awarded

Dr. Charles Frederick Burgess, inventor of the standard Burgess battery, was awarded the Edward Goodrich Acheson Medal and $1000 prize at the eighty-second convention of the Electrochemical Society in Detroit on October 8. This award was established some years ago by Dr. Acheson, inventor of artificial graphite and carborundum. Dr. Burgess, the seventh recipient, is the inventor of numerous diversified products and processes, including the dry cell, electrolytic iron, and sterilization with chlorine. He is a director of the Burgess Battery Company, Burgess Cellulose Company, Burgess-Parr Company, and the Thordarsen Electric Manufacturing Company.

★

### Symposium on Ultra-Short Electromagnetic Waves

Remarkable advances in the field of ultra-high frequency radio equipment have been made due to the extensive research and development work that has been carried on during the past few years. To keep abreast with the times a symposium on ultra-short electromagnetic waves will be presented by the Basic Science Group of the New York Section, A.I.E.E. This will consist of six lectures to be held at 7:00 P.M. in Room 301, Pupin Hall, Columbia University. Subjects and speakers will be:

**Electromagnetic Theory** by PROFESSOR E. WEBER, *Polytechnic Institute of Brooklyn*, November 25.
**Transmission Line Theory** by DR. J. R. RAGAZZINI, *Columbia University*, December 16.
**Generation** by DR. I. E. MOUROMTSEFF, *Westinghouse Electric and Manufacturing Company*, January 20.
**Guided Propagation** by DR. S. A. SCHELKUNOFF, *Bell Telephone Laboratories*, February 10.
**Radiation** by MR. A. ALFORD, *International Telephone and Radio Manufacturing Company*, March 17.
**Reception** by MR. BERTRAM TREVOR, *RCA Communications, Inc.*, April 21.

Further information on this symposium may be obtained from Mr. H. E. Farrer, A.I.E.E. Headquarters, 33 West 39 Street, New York, New York.

★

### Salvage Suggestions for Industry

The American Industries Salvage Committee, under the chairmanship of Robert W. Wolcott, President of the Lukens Steel Company, offers the following ten suggestions for organizing the collection of industrial scrap:

1. Head a campaign by appointing an able executive of the company, armed with authority to act.
2. Use posters, illustrations, pay-envelope stuffers, and all other publicity means to enlist all employees in your scrap campaign.
3. Comb plant and yards for dormant scrap and unusable and abandoned equipment.
4. Survey all plant equipment, particularly idle, standing, or discarded machines.
5. Classify and segregate scrap and supervise its handling.
6. Make each foreman responsible for preventing spoilage and waste in his department.

7. Report promptly equipment which is obsolete. If equipment has not been used in three months, and it cannot be proved that it will be used in the next three, turn it over where it can be used, or scrap it.

8. Salvage usable parts from equipment marked for scrapping.

9. Speed return of scrap to mills and refineries through existing channels. Report regularly on collections of scrap to the Industrial Salvage Committee set up by the War Production Board in your community.

10. Enforce monthly re-checks in every department to find scrap material previously overlooked.

★

### War on Fuel Waste

Information to assist the home owner in fuel conservation methods, the procedure of converting oil burner heating plants to coal, and the proper methods of firing coal and other fuels in heating equipment are outlined in a new pamphlet entitled *War on Fuel Waste—How to Do It*, just issued by the War Service Committee of the American Society of Heating and Ventilating Engineers. Reprints for distribution can be supplied for $8.00 per thousand. Single copies will be sent upon request by the American Society of Heating and Ventilating Engineers, 51 Madison Avenue, New York, New York, when accompanied by a self-addressed return envelope or five cents in stamps to cover addressing and mailing.

★

### Bulletin of Mathematical Biophysics

The December, 1942 issue of *The Bulletin of Mathematical Biophysics*, edited by N. Rashevsky and published by the University of Chicago Press, will contain the following papers:

**An Analysis of the Shapes of a Cell during Division with Particular Reference to the Role of Surface Tension.** H. D. LANDAHL

**Equilibrium Shapes in Non-Uniform Fields of Concentration.** H. D. LANDAHL

**Cellular Forms: The Tri-Axial Cell.** ALSTON S. HOUSEHOLDER

**The Linear Theory of Neuron Networks: The Static Problem.** WALTER PITTS

**Some Problems in Mathematical Biophysics of Visual Perception and Aesthetics.** N. RASHEVSKY

★

### New Laboratories Dedicated

On September 27 "Electron House," new home of RCA Laboratories, was dedicated at Princeton, New Jersey. The building is a 488-foot long, three-story structure with 150 laboratory bays in which almost every kind of research related to radio and electronics can be undertaken. A high frequency laboratory is located on the roof, and there is a field laboratory in one corner of the 260-acre area. For research with fluorescent or "glow" materials, the chemical laboratory has several dust-proof rooms. The rooms of the optics laboratory can be opened up to form a space of several hundred feet in which to test long beams of light. There is also a free sound room, three stories high, heavily padded and acoustically "dead."

Supply shafts of unique design run from basement to roof in 104 locations to carry electricity, gas, water, and compressed air and gases to 420 work benches. There are also little workshops in the corridors on each floor where a research man can make for himself a part which he needs immediately.

★

### Factory Expansion

Another large expansion in factory space and production facilities has been made by the Ohmite Manufacturing Company, Chicago, Illinois. This is in addition to the big increase in plant size announced in October, 1941. Essential space and facilities are provided for the production of units to meet the greatly increased demand occasioned by the war effort.

★

### Necrology

Dr. Arnold Berliner, founder and editor of *Die Naturwissenschaften*, died on March 23, 1942.

Ernest Calvin Bryant, Professor Emeritus of Physics at Middlebury College, died on September 7 at the age of seventy-five.

Raymond H. Danforth, Professor of Mechanical and Hydraulic Engineering and Head of Department at the Case School of Applied Science, died on August 31 at the age of sixty-four.

Joseph W. Gavett, Jr., Professor of Mechanical Engineering and Chairman of Department at the University of Rochester, died on August 28 at the age of fifty-three.

Daniel A. Lehman, Professor Emeritus of Mathematics and Astronomy at Goshen College, died on September 8 at the age of eighty-two.

Arthur Laidlaw Selby, Emeritus Professor of Physics at University College, Cardiff, Wales, died on July 22 at the age of eighty-one.

Arthur H. Thomas, President of the Arthur H. Thomas Company, dealers in laboratory apparatus and reagents, died on August 31.

★

### Calendar of Meetings

**November**

7 American Physical Society, New York State Section, Albany, New York
16–18 American Institute of Chemical Engineers, Cincinnati, Ohio
27–28 American Physical Society, Chicago, Illinois
27–28 American Mathematical Society, Notre Dame, Indiana
30–Dec. 4 American Society of Mechanical Engineers, New York, New York

**December**

27–Jan. 1 American Society of Naturalists, New York, New York
28–31 American Physical Society, New York, New York
28–Jan. 2 American Association for the Advancement of Science, New York, New York
28–Jan. 2 American Standards Association, New York, New York

# Contributed Original Research

## A Rapid Infra-Red Analytical Method for Hydrocarbon Mixtures and A Routine Spectrophotometer for Plant Control*

R. Robert Brattain and O. Beeck
*Shell Development Company, Emeryville, California*
(Received July 21, 1942)

A rapid method has been developed for the infra-red spectroscopic analysis of simple mixtures of hydrocarbons. Through the construction of special equipment this method has been reduced to a routine procedure which can be carried out by comparatively inexperienced operators. The instruments built have proven sufficiently sturdy to be used for control work in plant operation. By basing the analytical method on an empirical calibration, the calculations involved have been reduced to a minimum. The instruments are relatively simple and permit the use of comparatively large spectral slit widths. This has eliminated the necessity for amplification of the thermopile output energy and for the determination of any continuous absorption curve. However, small spectral slit widths can be attained with the more versatile spectrophotometer if this should be required.

### INTRODUCTION

ADVANCES in the petroleum industry during the last decade and a half have resulted in the manufacture of many diversified products from petroleum hydrocarbons. These advances have greatly increased the need for analysis of hydrocarbon mixtures and particularly of mixtures of isomers. New analytical methods in this field are not only essential to research but are equally necessary for the control of large scale plants. Plant control calls for simple and sturdy equipment as well as for rapidity of analysis. An analytical method, results of which lag a day behind sampling a plant which produces thousands of gallons of product per day, is not really effective regardless of its accuracy. The most commonly used method, distillation, requires from five to twenty hours depending on the accuracy demanded and the composition of the sample.

Differences in simple physical properties such as density or refractive index can be used for binary mixtures, except for isomers where such differences are usually too insignificant. Different hydrocarbons show, however, differences in their absorption spectra in the region from $2\mu$ to $25\mu$, where absorption is due to molecular vibrations. Isomers show relatively large differences between their spectra.

The possibilities for infra-red spectroscopy as an analytical tool have long been recognized. However, due mainly to experimental difficulties, infra-red spectroscopy has been used largely as a qualitative method only. Numerous instances where infra-red absorption spectra have been used for the identification of a compound or for establishing the presence of a particular chemical bond can be found in the literature. Recently, under the stimulus of an expanding chemical industry, several efforts have been made to use infra-red absorption spectra as the basis for rapid quantitative analysis, although only one report[1] of this nature is available in the literature. The present paper will discuss in detail the analytical procedure as exemplified by isomer analysis and will describe a versatile routine instrument suitable for plant control and for the analytical laboratory.

Greatest simplicity and rapidity was achieved only when the analytical procedure was based on a detailed study of the specific mixture to be analyzed. This does not mean that the whole

* Presented before the American Physical Society, Pasadena, California, June 1941. Phys. Rev. **60**, 161 (1941).

[1] Norman Wright, "Application of infra-red spectroscopy to industrial research," Ind. Eng. Chem. Anal. Ed. **13**, 1 (1941).

spectrum has to be studied in much detail except for the narrow regions from which final wave-lengths or prism positions are to be selected. Experience has taught that analysis based on absorption measurements at one or several fixed prism positions is superior to that obtained from a complete mapping of the spectral regions in question. This is particularly true when spectroscopic analysis is to be used in plant control, where a great number of samples of similar composition are to be analyzed.

A spectroscopic analytical method, which is to be used for routine analysis and plant control, should meet the following general requirements. The method should:

1. Be as rapid as is consistent with the precision required.
2. Yield the actual analysis from the experimental readings with a minimum of calculation.
3. Require only a simple procedure performed with a consistent and sturdy apparatus.

The restrictions imposed by these requirements on analytical procedure and instrument will be illustrated by a specific example. This illustration will show that some of these restrictions apply to the routine analysis of other mixtures.

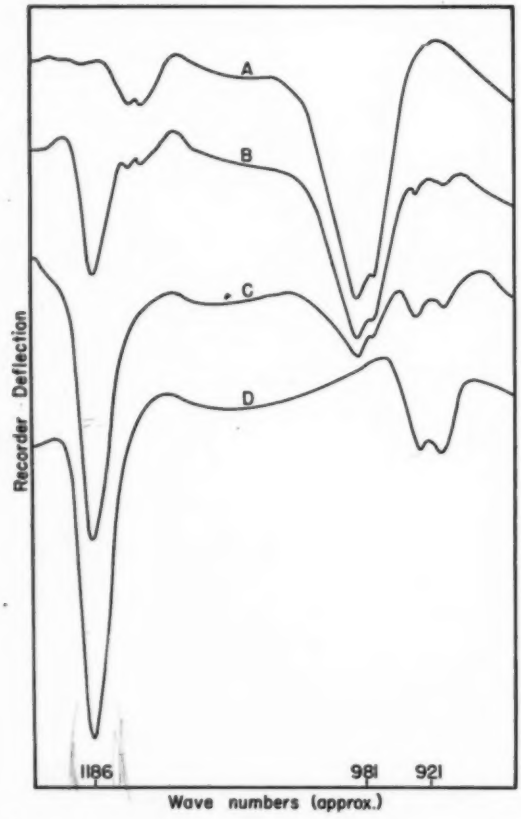



FIG. 1. Spectra of four samples containing different proportions of the butane isomers. *A* 99.4 percent *n*-butane; *B* 30 percent isobutane; *C* 70 percent isobutane; *D* 99.7 percent isobutane.

## ANALYSIS OF *n*-BUTANE-ISOBUTANE MIXTURES

For a two-component sample, the above general requirements are fulfilled most easily by a system requiring only the determination of the percent transmission at a single spectral position. With such a system, the largest change in percent transmission per unit change in isobutane content is achieved by choosing the proper values of the three variables: spectral position, spectral slit width, and specific absorption path. The optimum values of these variables may change with changing isobutane content of the samples under investigation. However, a rapid routine analysis procedure must handle all samples with a minimum of adjustments for sample composition.

The spectra between the wave numbers 1250 and 870 of four different mixtures of the butane isomers are shown in Fig. 1. It is seen that large differences exist in this region between the spectra of these isomers.

The absorption band of isobutane at 1186 $cm^{-1}$ was chosen as the spectral position. This band is freest from interference due to *n*-butane and is closer to the energy maximum of the source than the other bands shown in Fig. 1. The absorption maximum of this band is used as the exact spectral position in the case of butane mixtures. However, in the case of another mixture a position on the side of an absorption band may give a larger change in percent transmission per unit change of the unknown component. This situation could arise if interfering absorption bands are present.

The spectral slit width chosen for the analysis of isobutane in *n*-butane was about 65 $cm^{-1}$. This value was selected to give the maximum energy strongly absorbed by isobutane in this spectral region.

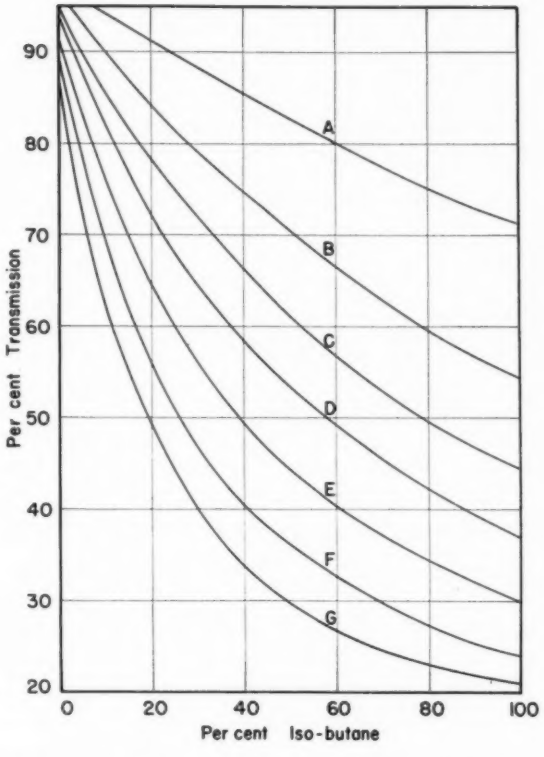



FIG. 2. Percent transmission plotted against percent isobutane for various pressures (in mm of Hg) of sample in a 15-cm absorption path. *A* 50 mm, *B* 100 mm, *C* 150 mm, *D* 200 mm, *E* 275 mm, *F* 375 mm, *G* 475 mm.

Curves showing percent transmission plotted against percent isobutane for various gas pressures of the sample in a 15-cm absorption cell are given in Fig. 2. From these data, curve *E* (275-mm pressure) was chosen as the most suitable over the entire 0 to 100 percent isobutane range and curve *C* (150-mm pressure) was chosen as the best if the sample contains more than 70 percent isobutane. Samples whose compositions were accurately known were used to obtain these curves.

To summarize, the routine infra-red analysis of *n*-butane-isobutane mixtures is made as follows:

The energy traversing the instrument and composed of the spectral interval from 1217 to 1153 $cm^{-1}$, is read both with and without the sample in the light path. From these readings, the percentage of energy transmitted by the sample is calculated. The isobutane content (in percent) is then read from a calibration curve. The calibration curve is plotted from readings taken on synthetic samples whose compositions are accurately known.

This system has been used, at this laboratory, to analyze 5500 samples in the past 22 months. In addition, two special instruments using this analysis procedure are now being used for plant control work. The accuracy of these routine analyses is about ±0.5 percent of the total sample. The accuracy is slightly better for samples low in isobutane and slightly worse if the isobutane content is above 85 percent. The time required for analysis is 20 minutes per sample. This figure includes time spent turning the instrument on and off, preparing the samples, calculating the results, periodically checking the method, and making out analysis reports. It was arrived at by dividing the total time spent on these analyses by the total number (5500) of samples analyzed.

The time is 15 minutes or less per sample if the number of analyses is sufficient to require 8 hours continuous work by an experienced operator. The average time per sample was increased to 20 minutes because six operators were trained within the 22 months with the spectrophotometer seldom in operation for more

TABLE I.

| Date analyzed | % Isobutane by infra-red | % Isobutane synthetic | % Isobutane duplicate | Low temperature dist. % Isobutane | % *n*-Butane | % Propane | % Pentane |
|---|---|---|---|---|---|---|---|
| 10/ 4/40 | 10.0 | 9.9 | | | | | |
| 10/ 9/40 | 45.8 | 45.6 | | | | | |
| 10/11/40 | 30.0 | 30.17 | | | | | |
| 10/11/40 | 45.9 | 45.6 | | | | | |
| 10/11/40 | 100.5 | 99.7 | | | | | |
| 6/11/41 | 20.1 | 19.86 | | | | | |
| 6/11/41 | 19.8 | 19.86 | | | | | |
| 6/11/41 | 29.8 | 30.17 | | | | | |
| 6/11/41 | 99.1 | 99.7 | | | | | |
| 7/14/41 | 20.0 | 20.12 | | | | | |
| 7/14/41 | 30.0 | 30.08 | | | | | |
| 7/14/41 | 40.0 | 40.08 | | | | | |
| 7/14/41 | 80.5 | 79.79 | | | | | |
| 10/14/41 | 43.5 | 43.9 | | | | | |
| 5/ 8/40 | 23.25 | | | 22.6 | 75.6 | 0.0 | 1.8 |
| | 12.0 | | | 11.2 | 88.4 | 0.0 | 0.4 |
| | 54.9 | | | 54.9 | 43.7 | 0.2 | 1.2 |
| 5/16/40 | 36.9 | | | 36.5 | 62.8 | 0.0 | 0.7 |
| | 16.0 | | | 16.3 | 83.3 | 0.0 | 0.4 |
| | 61.0 | | | 60.1 | 36.3 | 0.7 | 2.9 |
| 6/31/40 | 37.2 | | | 36.5 | 62.5 | 0.1 | 0.9 |
| 7/22/40 | 37.4 | | | 37.9 | 61.2 | 0.2 | 0.7 |
| 6/ 3/40 | 27.4 | | 27.7 | | | | |
| 6/10/40 | 48.4 | | 48.6 | | | | |
| 6/18/40 | 47.0 | | 47.5 | | | | |
| 6/11/40 | 22.2 | | 21.8 | | | | |

than 4 hours per day. One operator analyzed 38 samples in 6 hours, including calculations and other operations incidental to the actual analyses.

## ACCURACY CHECKS AND INTERFERENCE FROM OTHER SUBSTANCES

The results of some of the checks of this infra-red method of analyzing butane mixtures are presented in Table I. Three different bases of comparison are given:

1. Synthetic mixtures whose compositions are accurately known.
2. Values obtained by low temperature distillation. The complete composition is given for these samples.
3. Duplicate spectroscopic values taken at least one day apart.

The accuracy of a low temperature distillation analysis is about $\pm 0.5$ percent.

Propane, *n*-pentane, and isopentane are the impurities most likely to be present in a *n*-butane-isobutane sample. All of these compounds have some absorption in the region of the 1186 $cm^{-1}$ isobutane absorption band. However, small quantities of these substances can be present without impairing the accuracy (0.5 percent) of the above simple analytical method. The maximum allowable amounts are 10 percent propane, 5 percent *n*-pentane, and 2.5 percent isopentane. Larger amounts of these impurities could probably be corrected for by readings at other wave-lengths. This would be possible on the more versatile routine spectrophotometer described below.

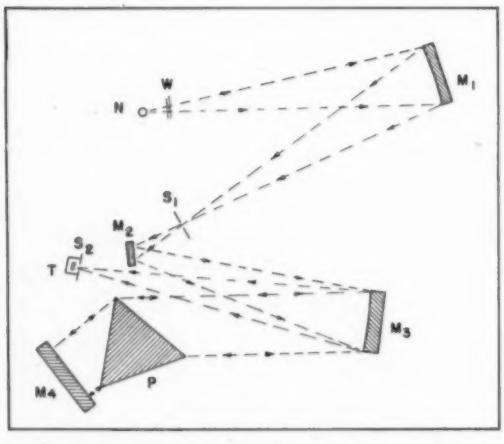



FIG. 3. The light path diagram of the routine infra-red spectrophotometer.

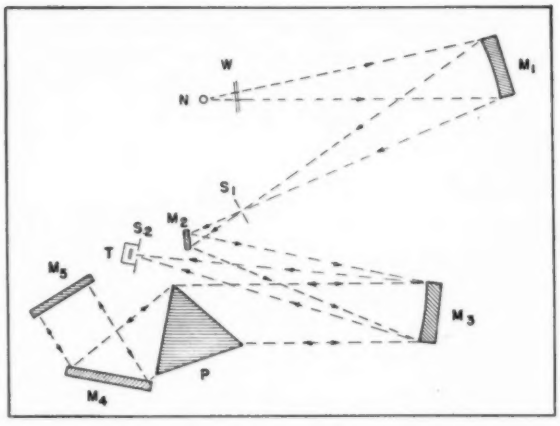



FIG. 4. The light path diagram of the versatile routine infra-red spectrophotometer.

None of the unsaturated four-carbon compounds, butenes, and butadiene, has large enough absorption at 1186 $cm^{-1}$ to cause small amounts of these compounds to interfere with this method of analyzing for isobutane.

## ANALYSIS OF MIXTURES OF BUTENE-1 AND BUTENE-2

The same system of analysis has been applied to mixtures of butene-1 and butene-2. The values of the three variables used for this analysis are:

1. Spectral position—absorption maximum of the butene-1 band at 1841 $cm^{-1}$.
2. Spectral slit width—about 65 $cm^{-1}$.
3. Specific absorption path—400-mm pressure of gas in a 15-cm cell.

About 250 such samples have been analyzed by this infra-red method. The accuracy and time per sample are the same as for butane mixtures.

## INSTRUMENTS

Two spectrophotometers using the same auxiliary equipment have been developed for the routine analysis of hydrocarbon mixtures. The first is designed for procedures in which it is necessary to determine transmission at only one spectral position. The second instrument is an extension of the first and allows measurements at several preadjustable spectral positions. This latter instrument will therefore serve several

single-position analyses as well as more complicated analyses requiring readings at several spectral positions for one analysis.

## SIMPLE SPECTROPHOTOMETER

The light path of the first spectrophotometer is shown in Fig. 3. The aperature of the instrument is $f$:5. $M_2$ and $M_4$ are plane mirrors, $M_1$ is a spherical mirror, and $M_3$ is a parabolic mirror ground 14° off axis. Its diameter is 4 cm and its focal length 20 cm. All mirrors are glass with a reflecting (front) surface of evaporated aluminum. The 60° rocksalt prism $P$ has refracting faces 4 cm high and 6 cm long. The thermopile $T$, immediately behind the second slit $S_2$, has four receiving and four compensating junctions and is evacuated. An absorption cell for gas samples is attached to the first slit $S_1$ on the side toward $M_1$. Cells for liquid samples could be inserted either between the source $N$, and the entrance window $W$, or between $W$ and $M_1$. The source is a special Nernst glower lamp and will be described in detail later. A shutter and a diaphragm for regulating the energy entering the instrument are located between the source and the entrance window. The rocksalt entrance window $W$ is coated with MgO powder to filter out visible radiation. The prism, the cell windows, and the entrance window are protected from normal atmospheric water vapor with a thin coat of polystyrene. This is of use mainly during adjustment as the air inside the gas-tight case is dried during routine operation. The slits $S_1$ and $S_2$ are simple knife edges which can be clamped after adjustment. All parts are clamped to a heavy cast-iron base plate and, with the exception of the source, are enclosed in a gas-tight steel case lined with insulating material.

The prism and Littrow mirror $M_4$ are adjusted to minimum deviation for the wave-length which is used for analysis and clamped in this position. This is accomplished by using the absorption band, on which transmission measurements are to be made, as an index point after having first adjusted the prism to minimum deviation for some visible wave-length. The parabolic mirror $M_3$ is the limiting aperture of the spectrophotometer. This fact in addition to the MgO coating on the entrance window, the blackening of all non-optical parts, and careful shielding eliminates false energy. No detectable false energy reaches the thermopile $T$, with the prism set for the 1186-cm$^{-1}$ isobutane band. At this setting the energy striking the thermopile gives a deflection of 200 mm at 1-m with the thermopile connected to a galvanometer of 0.1-$\mu$v/mm sensitivity.

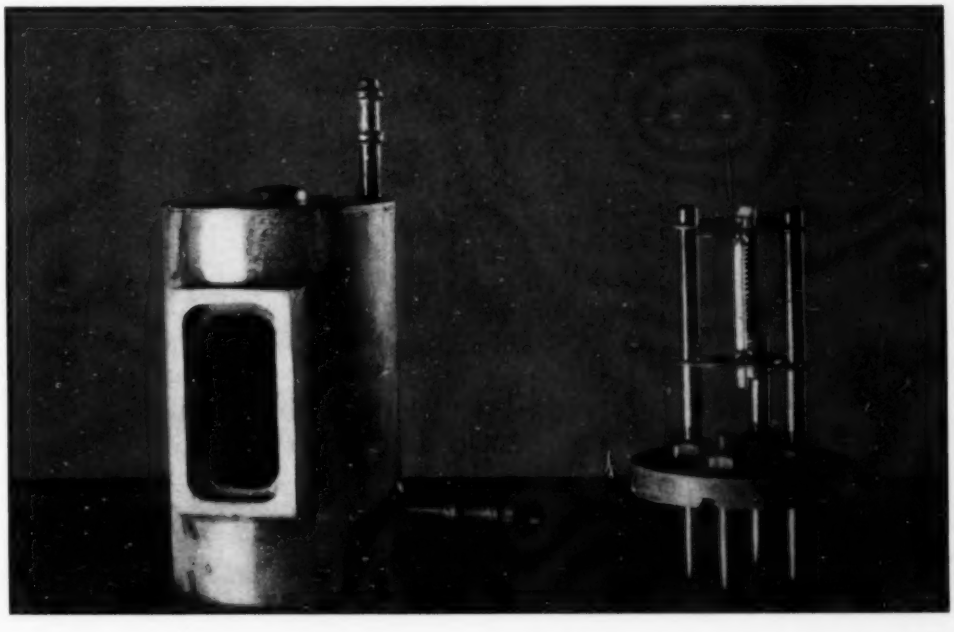


FIG. 5. The Nernst glower radiation source. The water jacket is shown on the left and the internal assembly on the right.

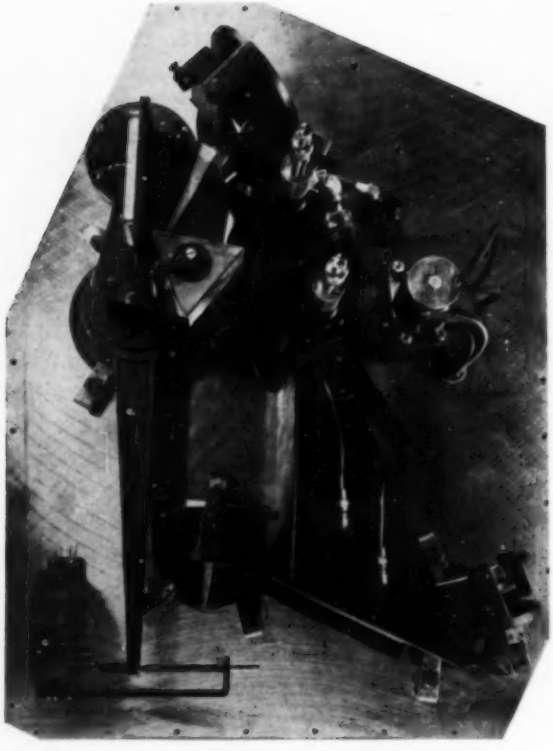


FIG. 6. Versatile analytical infra-red spectrophotometer assembled on base plate.

### VERSATILE SPECTROPHOTOMETER[2]

Figure 4 is the light path of the more versatile spectrophotometer for routine analysis.[2] The principal differences between this instrument and the one described above are in the slits and the prism mounting. The prism $P_1$ and the Wadsworth mirror $M_4$ are mounted on a turntable. The angular position of this table and, consequently, the wave-length of the radiation arriving at the thermopile $T$ are controlled by a tangent arm resting against stops. There are eighteen such stops which are adjustable from outside the spectrophotometer case. Since one of these is used for checking the focus at the sodium $D$ line, this allows the instrument to be set rapidly and reproducibly to determine the transmission of a sample at seventeen spectral positions. This instrument and the simpler spectrophotometer were designed for work in the $2\mu$- to $15\mu$-region. However, other spectral regions can be covered by using a material other than rocksalt for the prism and windows.

Since the different wave-length settings of the prism require different slit widths and since more complicated analytical procedures may require smaller spectral slit widths, slits $S_1$ and $S_2$ are adjustable from outside of the instrument case. Also, slit $S_1$ is curved sufficiently to compensate for the curvature of the image caused by the prism. Furthermore, the change of the refractive index of the prism, with temperature, causes a shift in the wave-length of the radiation reaching the thermopile, and when small spectral slit widths are used, this shift will effect the accuracy of analysis. Consequently, the holder for the Littrow mirror $M_5$ is built so that it rotates $M_5$ to correct for the change of the refractive index of the prism. This refinement is not necessary with the large spectral slit widths used in the simpler spectrophotometer described first.

### RADIATION SOURCE

The parts of the special Nernst glower source are shown in Fig. 5. The water jacket (6 cm high) on the left is made from standard sizes of brass tubing. If the source is operated in a fire-restricted area, a rocksalt window is sealed into the recess in the flange. Under these conditions, the screw in the top can be removed periodically to replenish the air (oxygen) inside the jacket. The right-hand view shows the Nernst glower and its starting heater assembled. The glower and the heater are located, with respect to the stainless steel studs, by the two mica washers. The studs are spaced in the heat-resistant insulating disk so that they fit a standard four-pronged base. This disk for holding the studs also serves as the bottom end of the jacket. The assembled lamp on its base block is shown in the general view of a spectrophotometer given in Fig. 6. All holes which affect the position of the Nernst glower with respect to the base block are jig drilled, allowing the source lamp to be replaced without readjusting any other parts of the spectrophotometer.

The ballast resistance, the transformer for the heater, and a relay are in a separate box. The circuit is such that closing the switch to 110-volts a.c. impresses the proper voltage across

[2] Now being manufactured by the National Technical Laboratories, South Pasadena, California.

both the heater and the Nernst glower. As soon as the glower reaches the temperature at which it conducts appreciable current the relay shuts off the heater.

Due to the high operating temperature certain precautions in the construction of the internal parts are necessary. All studs, spacers, and nuts are made of stainless steel. The terminals are connected to the Nernst glower by the method described by Ebers and Nielsen[3] and the special heat-resistant wire, Kanthol A, is used for the heater.

## AUXILIARY EQUIPMENT AND GENERAL FEATURES

The view of the versatile routine infra-red spectrophotometer shown in Fig. 6 serves to illustrate the type of construction used in both instruments. The heavy cast base plate for this instrument is 16 inches by 22 inches and the gas-proof, heat-insulating cover is 7 inches high (see Fig. 7). The simpler spectrophotometer fits on a base plate 15 inches by 18 inches. The holders for the various optical parts are sufficiently sturdy so that the instrument has remained in adjustment when shipped great distances.

A complete analysis unit, except for the galvanometer, is shown in Fig. 7. The gas-handling system, on the right, is used to fill the absorption cell with a standard pressure of sample. Time is saved by sweeping the system with air between samples. The old sample plus the sweeping air is removed by an auxiliary vacuum line and the Hyvac pump is used only to remove the remaining air. All controls including the shutter valve, the galvanometer scale, and the source switch, are arranged so that they can be manipulated from the same position as the gas-handling system. The circuit for controlling the source is shown under the table in a separate gas-tight case. This gas-tight feature is extended to the wiring so that the unit can be operated in fire-restricted areas.

[3] E. S. Ebers and H. H. Nielsen, "A method for increasing the life of Nernst glowers," Rev. Sci. Inst. **11**, 429 (1940).

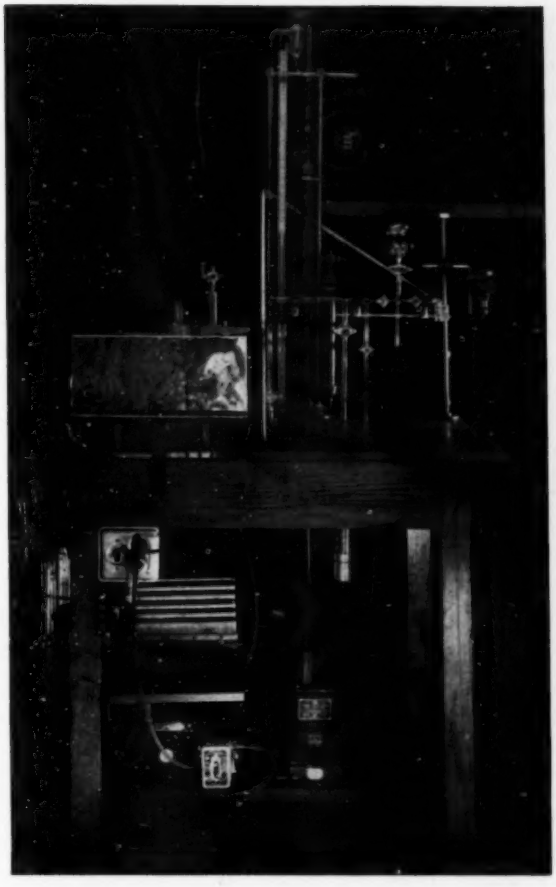


FIG. 7. Versatile analytical infra-red unit including the spectrophotometer, the gas-handling unit, and auxiliary equipment.

The pumping system, on the end of the table, is used to evacuate the thermopile. However, activated charcoal traps, evacuated and sealed at high temperatures, have been used to maintain the thermopile vacuum in other instruments.


## ACKNOWLEDGMENTS

The authors wish to express their indebtedness to Mr. B. M. Burchard for valuable assistance in connection with mechanical design, particularly with regard to the special source lamp. Mr. H. J. Fanger and Mr. H. Gruenhagen of Fanger Research and Manufacturing Company have also contributed many helpful suggestions on the construction of the mechanical parts of the spectrophotometers.

# Steady State Currents of Electrical Networks

D. L. WAIDELICH
*University of Missouri, Columbia, Missouri*
(Received August 3, 1942)

It is shown that the methods of operational circuit analysis may be extended to give the steady state current of a circuit not only in the well-known Fourier series form but also in the form of the sum function of this Fourier series. This sum function is very useful in determining the wave form of the current. Three methods of obtaining the steady state current are given along with their restrictions, and these methods involve a real integral, a finite series, and a complex integral. An example of the use of these methods is also given.

## INTRODUCTION

IN engineering applications such as welding, rectification, and television, periodic non-sinusoidal voltages and the resulting steady state currents have been encountered more and more frequently. The Fourier series method of solution for these currents has been used almost to the exclusion of all other methods, because it offers a straightforward method of approach. On the other hand, the current wave form is both tedious and difficult to obtain by this method. The problem of obtaining a finite or "closed" expression for the steady state currents* rather than an infinite Fourier series has led to the realization that the equation of these currents may be obtained over one period, and this equation may be called the sum of the Fourier series. Carnahan[1] while working with television problems has obtained the steady state currents resulting from the application of simple periodic voltages to networks. Hamlin[2] has generalized the work of Carnahan and gives an equation for the steady state current applicable to all periodic voltages and resembling somewhat Heaviside's expansion theorem. This paper will concern itself with developing various integral expressions for these currents. In addition a new method of obtaining this current involving an integral in the complex plane somewhat like the Bromwich-Wagner integral will be given. All of the methods discussed here except the complex integral method will give either the Fourier series form of the current or the sum of this Fourier series depending in general on whether the applied voltage is expressed as a Fourier series or as the sum of a Fourier series. Both of these types of solutions may be obtained from the complex integral also by using different methods of evaluating the integral.

* *Note added in proof.*—One method that has been used in calculating the steady state current as a sum function is that discussed by W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, New York, 1939), first edition, p. 377. The disadvantages of this method as compared to those discussed in the present paper are that (1) the boundary conditions must be known and (2) the values of the arbitrary constants must be obtained as the solutions of a group of simultaneous equations.

[1] C. W. Carnahan, Proc. I. R. E. **23**, 1393 (1935).
[2] E. W. Hamlin, J. Frank. Inst. **233**, 257 (1942).

## THE STEADY STATE CURRENT EXPRESSED BY INTEGRALS

The operational impedance of the circuit to which the periodic voltage is applied is $Z(p)$, and the reciprocal of this impedance $1/Z(p)$ is assumed to have as its singularities poles of finite order all located on or to the left of the imaginary axis. This is true for all passive electrical networks with lumped linear elements. In addition the circuit is assumed to be not impulsive in nature. The voltage applied to the circuit $e(t)$ has a period $T=2\pi/\omega$, and may be expressed as a Fourier series in terms of the time $t$:

$$e(t)=\sum_{n=-\infty}^{+\infty} E_n\epsilon^{jn\omega t}, \tag{1}$$

where

$$E_n=\frac{1}{T}\int_0^T \epsilon^{-jn\omega t}e(t)dt \text{ and } n \text{ is an integer.}$$

The Laplacian transform of this voltage is

$$E(p)=\int_0^\infty \epsilon^{-pt}e(t)dt = \sum_{n=-\infty}^{+\infty}\frac{E_n}{p-jn\omega}. \tag{2}$$

The voltage transform has all of its singularities on the imaginary axis equally spaced at intervals of $\omega=2\pi/T$, and, in addition, all of the singularities are poles of order one. None of the singularities of $1/Z(p)$ and $E(p)$ may coincide, because if they did, an infinite steady state current would result. This is the case of a network without resistance and resonant to any one of the fundamental or harmonic frequencies of the applied voltage.

Duhamel's theorem states that if $e(t)$ is the voltage applied to a circuit at $t=0$ and $A(t)$ is the indicial admittance of the circuit, the current $i(t)$ is given by:

$$i(t)=A(0)e(t)+\int_0^t e(t-\tau)A'(\tau)d\tau. \qquad (3)$$

If $e(t)$ has a period $T$ and the time origin is shifted by means of the transformation $t=t'+nT$ where $n$ is a positive integer, (3) becomes:

$$i(t'+nT)=A(0)e(t'+nT)+\int_0^{(t'+nT)} e(t'+nT-\tau)A'(\tau)d\tau. \qquad (4)$$

As $n$ increases without bound, $i(t'+nT)$ approaches the steady state current $i_s(t')$. Dropping the primes and assuming that the integral converges, (4) is in the limit the steady state current:

$$i_s(t)=A(0)e(t)+\int_0^\infty e(t-\tau)A'(\tau)d\tau. \qquad (5)$$

This will always converge providing that the circuit is not impulsive ($1/Z(p)$ has no pole at infinity) and that all the poles of $1/Z(p)$ are to the left of the imaginary axis. The transient current $i_{TR}(t)$ is given by the difference between (3) and (5),

$$i_{TR}(t)=i(t)-i_s(t)=-\int_t^\infty e(t-\tau)A'(\tau)d\tau \qquad (6)$$

and converges for the same conditions as (5). Another way of obtaining these same results is to rewrite (3) as:

$$i(t)=A(0)e(t)+\int_0^\infty e(t-\tau)A'(\tau)d\tau-\int_t^\infty e(t-\tau)A'(\tau)d\tau. \qquad (7)$$

Since the third term on the right-hand side of (7) approaches zero as $t$ becomes very large, while the first two terms have a period $T$, the third term must represent the transient current $i_{TR}(t)$ as in (6), and the remaining two terms must be the steady state current $i_s(t)$ as in (5). There are six widely used forms of the Duhamel theorem, and for some of these forms there is the corresponding expression for the steady state current and for the transient current. These are listed in Table I along with the restrictions on their use. The first formula has the least number of restrictions on its use and probably is thus the most useful.

It is possible to transform the infinite integral of (5) into a finite integral:

$$i_s(t)=A(0)e(t)+\int_0^T e(t-\tau)f'(\tau)d\tau, \qquad (8)$$

where

$$f(t)=A(t)+A(t+T)+A(t+2T)+\cdots$$

and

$$f'(t)=A'(t)+A'(t+T)+A'(t+2T)+\cdots.$$

Put $\tau'=t-\tau+KT$ in (8) where $K$ is a positive or negative integer. Dropping the primes:

$$i_s(t)=A(0)e(t)+\int_{t+(K-1)T}^{t+KT} e(\tau)f'(t-\tau-KT)d\tau. \qquad (9)$$

When $K=0$,

$$i_s(t)=A(0)e(t)+\int_{t-T}^t e(\tau)f'(t-\tau)d\tau \qquad (10)$$

and this seems to be the most useful of all these integrals.

In the case of networks with no resistance, the poles of $1/Z(p)$ lie on the imaginary axis, and these steady state integrals may or may not converge. In every instance, however, the integrals may be evaluated by a summation process.[3] It must be remembered that no pole of $E(p)$ must coincide with one of $1/Z(p)$.

## THE STEADY STATE CURRENT AS A FINITE SERIES

The indicial admittance $A(t)$ may be expressed as a finite series if $1/Z(p)$ has no multiple poles

[3] V. Bush, *Operational Circuit Analysis* (John Wiley and Sons, New York, 1937), first edition, p. 186.

TABLE I. Expressions for the total, transient, and steady state currents.

| | Total current $i(t)$ | Transient current $i_{TR}(t)$ | Steady state current $i_s(t)$ Infinite integral |
|---|---|---|---|
| 1 | $A(0)e(t)+\int_0 e(t-\tau)A'(\tau)d\tau$ | $-\int_t^\infty e(t-\tau)A'(\tau)d\tau$ | $A(0)e(t)+\int_0^\infty e(t-\tau)A'(\tau)d\tau$ |
| 2 | $\frac{d}{dt}\int_0^t e(t-\tau)A(\tau)d\tau$ | $-\frac{d}{dt}\int_t^\infty e(t-\tau)A(\tau)d\tau$ | $\frac{d}{dt}\int_0^\infty e(t-\tau)A(\tau)d\tau$ |
| 3 | $e(0)A(t)+\int_0^t e'(t-\tau)A(\tau)d\tau$ | $e(0)A(t)-\int_t^\infty e'(t-\tau)A(\tau)d\tau$ | $\int_0^\infty e'(t-\tau)A(\tau)d\tau$ |

| | Steady state current $i_s(t)$ Finite integrals | | Restrictions on the use of $i_{TR}(t)$ and $i_s(t)$ integrals |
|---|---|---|---|
| 1 | $A(0)e(t)+\int_0^T e(t-\tau)f'(\tau)d\tau$ | $A(0)e(t)+\int_{t-T}^t e(\tau)f'(t-\tau)d\tau$ | 1. Circuit must not be impulsive.<br>2. All poles of $1/Z(\rho)$ must lie to the left of the imaginary axis.* |
| 2 | $\frac{d}{dt}\int_0^T e(t-\tau)f(\tau)d\tau$ | $\frac{d}{dt}\int_{t-T}^t e(\tau)f(t-\tau)d\tau$ | All of the above and<br>3. $A(\infty)=1/Z(0)$ must be zero. |
| 3 | $\int_0^T e'(t-\tau)f(\tau)d\tau$ | $\int_{t-T}^t e'(\tau)f(t-\tau)d\tau$ | All of the above and<br>4. $e(t)$ must be continuous. (No finite jumps.) |

* If the roots of $1/Z(p)$ lie on the imaginary axis, the integrals do not converge but may be evaluated by a process of summation.

(Heaviside expansion theorem):

$$A(t)=\frac{1}{Z(0)}+\sum_{m=1}^{m}\frac{\epsilon^{\lambda_m t}}{\lambda_m Z'(\lambda_m)} \tag{11}$$

and

$$A'(t)=\sum_{m=1}^{m}\frac{\epsilon^{\lambda_m t}}{Z'(\lambda_m)}, \tag{12}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_m$ are the poles of $1/Z(p)$. When (12) is substituted into (8),

$$i_s(t)=\frac{e(t)}{Z(\infty)}+\sum_{m=1}^{m}\frac{1}{(1-\epsilon^{\lambda_m T})Z'(\lambda_m)}\times\int_0^T\epsilon^{\lambda_m\tau}e(t-\tau)d\tau. \tag{13}$$

Similarly from (10):

$$i_s(t)=\frac{e(t)}{Z(\infty)}+\sum_{m=1}^{m}\frac{\epsilon^{\lambda_m t}}{(1-\epsilon^{\lambda_m T})Z'(\lambda_m)}\times\int_{t-T}^t\epsilon^{-\lambda_m\tau}e(\tau)d\tau. \tag{14}$$

Equation (14) gives the steady state current $i_s(t)$ directly in terms of the applied voltage $e(t)$ and the impedance function of the circuit $Z(p)$. The integrals of (14) may be evaluated graphically if necessary, although it is usually easier to fit an equation to the curve of $e(t)$. In a manner similar to this, it is possible to obtain the general equation of Hamlin[2] for the steady state current, and in fact his result is very similar but not nearly as compact as that of (13) or (14). These two expressions have the disadvantage, however, as compared to the other expressions given in this paper, that $1/Z(p)$ must have no multiple poles.

## THE COMPLEX INTEGRAL METHOD

From (1), the steady state current is

$$i_s(t)=\sum_{n=-\infty}^{+\infty}\frac{E_n}{Z(jn\omega)}\epsilon^{jn\omega t} \tag{15}$$

Each term of the series of (15) is equal to

$$\frac{E_n}{Z(jn\omega)}\epsilon^{jn\omega t}=\frac{1}{2\pi j}\int_{O_n}\frac{\epsilon^{pt}E(p)}{Z(p)}dp, \tag{16}$$

where $E(p)$ is defined by (2) and $O_n$ is a curve enclosing the point $p=jn\omega$ but no other poles of $E(p)$ or $1/Z(p)$. The steady state current would then be the sum of the integrals of (16) each taken around a curve $O_n$ enclosing the points $p=jn\omega$ where $n$ is an integer varying from $-\infty$ to $+\infty$. These points indicated in Fig. 1(*a*) are the singularities of $E(p)$ also, and in addition one

integral evaluated along a path $C$ enclosing all of the singularities of $E(p)$ and none of $1/Z(p)$ will give the same result as the sum of the integrals of (16). The steady state current is then

$$i_s(t)=\frac{1}{2\pi j}\int_C \frac{\epsilon^{pt}E(p)}{Z(p)}dp, \qquad (17)$$

where $C$ is the heavy line curve of Fig. 1($a$) and the direction of the path of integration is as indicated. The curve $C$ may be thought of as composed of two curves $C_1$ and $C_2$ of Fig. 1. If the circuit resistance is small or entirely absent, the poles of $1/Z(p)$ will lie very close to or on the imaginary axis. The curve $C$ along which the integral is taken must then be changed to a form such as that of Fig. 1($b$). The poles on the imaginary axis of $1/Z(p)$ are assumed to be $(ja)$ and $(-ja)$ where $a$ is real, and all the poles of

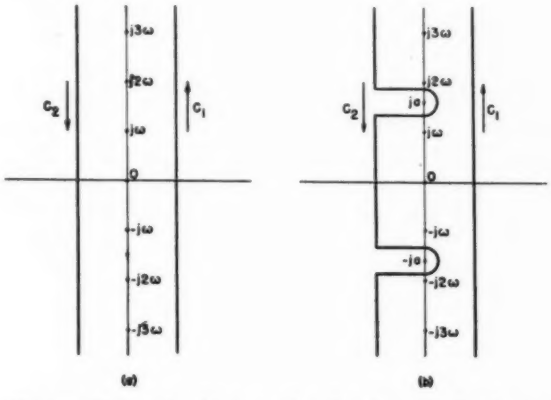



FIG. 1. The path of integration for the complex integral giving the steady state current as a Fourier series; ($a$) poles of $1/Z(p)$ not on the imaginary axis, ($b$) poles of $1/Z(p)$ on the imaginary axis.

$1/Z(p)$ must be to the left of $C_2$, so that (17) may give the steady state current. The singularities of $E(p)$ are the simple poles $(jn\omega)$, but it is not necessary that $E(p)$ have all these points as singularities. For example, if the periodic voltage $e(t)$ has no second harmonic voltage, the points $(j2\omega)$ and $(-j2\omega)$ are not singularities of the corresponding $E(p)$.

Once the transform $E(p)$ of the periodic voltage $e(t)$ and the impedance function $Z(p)$ are known, the steady state current may be determined by evaluating (17). In general this evaluation may be carried out by two general methods, one of which leads to the Fourier series and the other to the sum function of the Fourier series. To find the Fourier series expansion, Eq. (17) is evaluated by obtaining the sum of the residues at the poles inside curve $C$ of Fig. 1. This results in the series of (15) which is the steady state current.

The second method consists of rewriting (17) as a number of separate integrals and evaluating each one separately. Then

$$i_s(t)=\frac{1}{2\pi j}\int_C \frac{\epsilon^{pt}E(p)}{Z(p)}dp$$

or

$$i_s(t)=\frac{1}{2\pi j}\int_{C_1} \frac{\epsilon^{pt}E(p)}{Z(p)}dp-\frac{1}{2\pi j}\int_{C_3} \frac{\epsilon^{pt}E(p)}{Z(p)}dp, \qquad (18)$$

where $C_1$ and $C_3$ are the curves shown in Fig. 2($a$) along with their directions. Figure 2($b$) shows the same curves for the case in which $1/Z(p)$ has the poles, $p=\pm ja$. The curve $C_3$ is exactly the same as $C_2$ of Fig. 1 except that it is traversed in the opposite direction. All of the singularities of $E(p)$ must lie between $C_1$ and $C_3$, and the singularities of $1/Z(p)$ must be to the left of $C_3$. The first integral of (18) represents the total current $i(t)$ flowing[4] when a voltage $e(t)$ is applied at $t$ equal to zero to a circuit with impedance $Z(p)$. This integral

$$i(t)=\frac{1}{2\pi j}\int_{C_1} \frac{\epsilon^{pt}E(p)}{Z(p)}dp \qquad (19)$$

may be evaluated by a method to be outlined later. The second integral of (18) is the difference between the total current $i(t)$ and the steady state current $i_s(t)$. This current may be called the transient current $i_{TR}(t)$ and is given by

$$i_{TR}(t)=\frac{1}{2\pi j}\int_{C_3} \frac{\epsilon^{pt}E(p)}{Z(p)}dp. \qquad (20)$$

This transient current $i_{TR}(t)$ may be evaluated by finding the sum of the residues at the singularities of $1/Z(p)$ which must all lie to the left of $C_3$.

The voltage transform $E(p)$ may be written as:

$$E(p)=V_0(p)+\epsilon^{-pt}V_1(p)+\epsilon^{-p2T}V_2(p)+\cdots, \qquad (21)$$

where the $V$'s are functions of $p$ but contain no

[4] L. A. Pipes, J. App. Phys. **10**, 172 (1939).

$\epsilon^{-pT}$ terms. Since the $C_1$ contour is the same as the $Br_1$ contour of McLachlan,[5] Eq. (19) for $i(t)$ may then be expressed as a sum of integrals

$$i(t)=\frac{1}{2\pi j}\int_{C_1}\frac{\epsilon^{pt}V_0(p)}{Z(p)}dp + \frac{1}{2\pi j}\int_{C_1}\frac{\epsilon^{p(t-T)}V_1(p)}{Z(p)}dp+\cdots \quad (22)$$

and since all but the first integral of this series is zero[5] when $0<t<T$, the total current for the first period or when $0<t<T$, is

$$i(t)=\frac{1}{2\pi j}\int_{C_1}\frac{\epsilon^{pt}V_0(p)}{Z(p)}dp, \quad (23)$$

where this integral is evaluated by finding the sum of the residues at all the singularities of $1/Z(p)$ and $V_0(p)$. Finally the steady state current when $0<t<T$ is the difference of (23) and (20) or

$$i_s(t)=\frac{1}{2\pi j}\int_{C_1}\frac{\epsilon^{pt}V_0(p)}{Z(p)}dp - \frac{1}{2\pi j}\int_{C_3}\frac{\epsilon^{pt}E(p)}{Z(p)}dp. \quad (24)$$

This will give the sum function of the Fourier series, and since the steady state current $i_s(t)$ repeats each period, it is necessary to obtain the current only for one period. It is possible, however, to obtain the sum function for any period, if it is remembered that in (22) when $T<t<2T$, all but the first two integrals are zero; when $2T<t<3T$, all but the first three are zero, and so on.

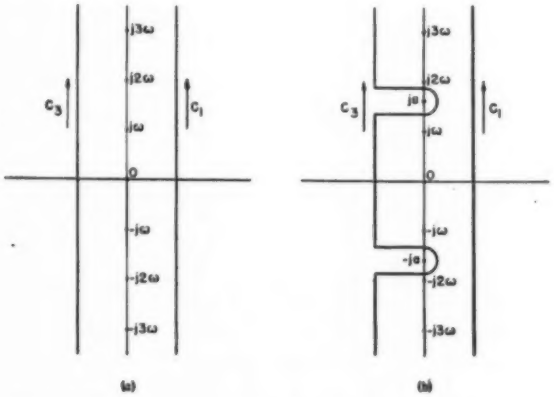



FIG. 2. The integration path used to obtain the steady state current as a sum function; (*a*) poles of $1/Z(p)$ not on the imaginary axis, (*b*) poles of $1/Z(p)$ on the imaginary axis.

[5] N. W. McLachlan, *Complex Variable and Operational Calculus* (Cambridge University Press, London, 1939), first edition, pp. 60 and 129.

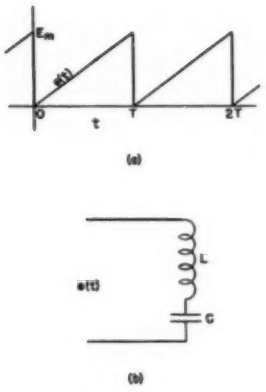



FIG. 3. The circuit and voltage used in the example; (*a*) applied voltage, (*b*) circuit.

In general the total current flowing when a periodic voltage $e(t)$ is applied to a circuit of impedance $Z(p)$ is by (19):

$$\begin{aligned} i(t)&=\frac{1}{2\pi j}\int_{C_1}\frac{\epsilon^{pt}E(p)}{Z(p)}dp \\ &=\sum \text{Residues at poles of } E(p) \\ &\quad +\sum \text{Residues at poles of } 1/Z(p). \end{aligned} \quad (25)$$

From (17), the steady state current is:

$$\begin{aligned} i_s(t)&=\frac{1}{2\pi j}\int_{C}\frac{\epsilon^{pt}E(p)}{Z(p)}dp \\ &=\sum \text{Residues at poles of } E(p). \end{aligned} \quad (26)$$

Consequently the transient current is by (20):

$$\begin{aligned} i_{TR}(t)=i(t)-i_s(t)&=\frac{1}{2\pi j}\int_{C_3}\frac{\epsilon^{pt}E(p)}{Z(p)}dp \\ &=\sum \text{Residues at poles of } 1/Z(p). \end{aligned} \quad (27)$$

The steady state current is then the difference of the expressions for the total current and for the transient current. In addition, since the total current may be evaluated in two different ways to produce two entirely different expressions, the

steady state current may be expressed by using these two expressions either as a Fourier series or as the sum function of the Fourier series.

### EXAMPLE

If the saw tooth voltage of Fig. 3(*a*) is applied to the series *L*-*C* circuit of Fig. 3(*b*), the steady state current may be calculated by the use of (14). If the voltage $e(t)$ is expressed as a Fourier series,

$$e(t)=\frac{E_m}{2T^2}-\sum_{n=1}^{\infty}\frac{E_m}{n\pi}\sin n\omega t, \tag{28}$$

the current will also be expressed as a Fourier series. The impedance of the circuit is $Z(p)=(1+p^2LC)/pC$ where $L$ is the inductance and $C$ is the capacitance. The poles of $1/Z(p)$ are $(ja)$ and $(-ja)$ where $a=1/(LC)^{\frac{1}{2}}$, and $1/(LC)^{\frac{1}{2}}$ must not equal $n\omega$, or resonance will result. Then

$$\begin{aligned}i_s(t)&=\frac{\epsilon^{jat}}{(1-\epsilon^{jaT})2L}\int_{t-T}^{t}\epsilon^{-ja\tau}\left[\frac{E_m}{2T^2}-\sum_{n=1}^{\infty}\frac{E_m}{n\pi}\sin n\omega\tau\right]d\tau\\&\quad+\frac{\epsilon^{-jat}}{(1-\epsilon^{-jaT})2L}\int_{t-T}^{t}\epsilon^{ja\tau}\left[\frac{E_m}{2T^2}-\sum_{n=1}^{\infty}\frac{E_m}{n\pi}\sin n\omega\tau\right]d\tau\\&=\sum_{n=1}^{\infty}\frac{E_m}{n\pi}\frac{\cos n\omega t}{(n\omega L-1/n\omega C)},\end{aligned} \tag{29}$$

where

$$\omega=2\pi/T.$$

If on the other hand the voltage is expressed as the sum function of the Fourier series, for $0<t<T$,

$$e(t)=\frac{E_m t}{T} \tag{30}$$

the steady state current by Eq. (14) is then, for $0<t<T$,

$$\begin{aligned}i_s(t)&=\frac{\epsilon^{jat}}{(1-\epsilon^{jaT})2L}\left[\int_{t-T}^{0}(\epsilon^{-ja\tau})\frac{E_m(\tau+T)}{T}d\tau+\int_{0}^{t}(\epsilon^{-ja\tau})\frac{E_m\tau}{T}d\tau\right]\\&\quad+\frac{\epsilon^{-jat}}{(1-\epsilon^{-jaT})2L}\left[\int_{t-T}^{0}(\epsilon^{ja\tau})\frac{E_m(\tau+T)}{T}d\tau+\int_{0}^{t}(\epsilon^{ja\tau})\frac{E_m\tau}{T}d\tau\right]\\&=\frac{CE_m}{T}\left\{1+\frac{aT}{2(1-\cos aT)}[\sin a(t-T)-\sin at]\right\},\end{aligned} \tag{31}$$

which is the sum function of the Fourier series for the current.

The steady state current may also be calculated by the use of the complex integral. For this the voltage transform is necessary and may be calculated[5] using (2),

$$\begin{aligned}E(p)&=\int_0^{\infty}\epsilon^{-pt}e(t)dt=\frac{1}{1-\epsilon^{-pT}}\int_0^{T}\epsilon^{-pt}e(t)dt\\&=\frac{E_m}{T(1-\epsilon^{-pT})p^2}[1-\epsilon^{-pT}(1+pT)].\end{aligned} \tag{32}$$

If the Fourier series is desired, Eq. (17) must be used. In this $C$ is a contour similar to that of Fig. 1(*b*) which includes the $E(p)$ poles, $p=jn\omega$ where $n$ is an integer, and excludes the $1/Z(p)$ poles, $p=\pm ja$. The steady state current is

$$\begin{aligned}i_s(t)&=\frac{1}{2\pi j}\int_C\frac{\epsilon^{pt}E_m[1-\epsilon^{-pT}(pT+1)]pC}{Tp^2(1-\epsilon^{-pT})(1+LCp^2)}dp\\&=\sum_{n=1}^{\infty}\frac{E_m}{n\pi}\frac{\cos n\omega t}{(n\omega L-1/n\omega C)}\end{aligned} \tag{33}$$

and this has been evaluated by summing the residues at each one of the poles, $p=jn\omega$. The result is the same as that of (29). If the sum function of the Fourier series is desired, Eq. (24) may be used. The $C_1$ and $C_3$ contours are similar to those of Fig. 2(*b*). Since

$$E(p)=\frac{E_m}{Tp^2}-\frac{E_m}{p}\sum_{r=1}^{\infty}\epsilon^{-prT} \tag{34}$$

the steady state current is, for $0<t<T$,

$$\begin{aligned} i_s(t)&=\frac{1}{2\pi j}\int_{C_1}\frac{\epsilon^{pt}E_m pC}{Tp^2(1+p^2LC)}dp \\ &\quad-\frac{1}{2\pi j}\int_{C_3}\frac{\epsilon^{pt}E_m[1-\epsilon^{-pT}(pT+1)]pC}{Tp^2(1-\epsilon^{-pT})(1+p^2LC)}dp \\ &=\frac{CE_m}{T}\left\{1+\frac{aT}{2(1-\cos aT)}\right. \\ &\quad\left.\times[\sin a(t-T)-\sin at]\right\}. \end{aligned} \tag{35}$$

The first integral is evaluated by summing the residues at the three poles, $p=0$, and $p=\pm ja$, while the second integral is evaluated only at the two poles, $p=\pm ja$. The result of (35) is again the same as that of (31).

### CONCLUSIONS

1. The steady state current of a circuit may be obtained from several different expressions which are given along with the restrictions on their use. These expressions consist of (a) various real integrals, (b) a finite series resembling Heaviside's expansion theorem, and (c) an integral in the complex plane.

2. The steady state current may be expressed either as a Fourier series or as the sum function of a Fourier series, both of which may be obtained by any of the methods given in 1. These methods are most useful in obtaining the sum function of the Fourier series but may also be used to obtain the Fourier series itself. The wave form of the current may be determined very easily from the sum function, but is very difficult to obtain from the series.

3. The transient current of a circuit is also given by a number of different equations.

---

## Formulas for the Inductance of Rectangular Tubular Conductors

THOMAS JAMES HIGGINS
*Illinois Institute of Technology, Chicago, Illinois*
(Received August 3, 1942)

Utilizing a function $H(x)$ defined by G. Stein, formulas are derived for the inductance per unit length of rectangular tubular conductors. An example illustrates use of these formulas. A table of Stein's function is given for values of the argument $x$ between 0 and 1 at intervals of 0.01.

THE mechanical and electrical advantages stemming from the use of rectangular tubular conductors have resulted in their extensive, and currently increasing, employment for feeding heavy current loads. To be noted, in particular, is their use for new bus installations supplying batteries of welders (in airplane, tank, and motor vehicle production) and banks of electric furnaces (for processing alloy and tool steels, aluminum, and other essential metals) in new plants necessitated by the defense program. In light of this accelerating employment it is obviously desirable to have formulas enabling the designer to calculate, quickly and accurately, the electrical constants of rectangular tubular conductors. For a specific conductor in a single or polyphase network, calculation of the resistance per unit length is easily effected; however, calculation of the inductance per unit length is usually more difficult.

For the network commonly used, i.e., a group of linear conductors of identical cross section, their axes parallel and coplanar, the corresponding sides of their right cross sections parallel, the writer has given a formula for calculating the inductance per unit length of a specific conductor

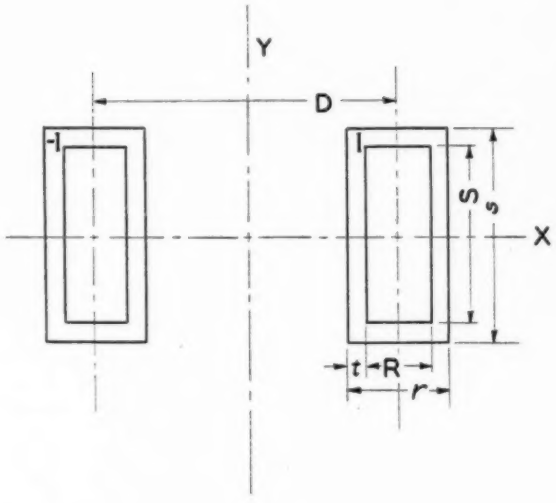



FIG. 1.

of the group. But though to this formula certain advantages accrue by virtue of its form—it is expressed as the algebraic sum of rapidly converging series, the required computation being readily performed on an automatic calculating machine—a certain unremovable disadvantage also recurs, i.e., the terms of the various series must be calculated to at least six decimal places, affording excellent possibilities for the inclusion of unsuspected numerical errors. Though precise checking and rechecking will, of course, reveal these errors, it is desirable to have an alternative and absolute means of detection: for instance, a second formula numerically independent of the first. *A formula of this kind is derived in Section 1; an example of its use is presented in Section 2.*

## 1. DERIVATION OF THE FORMULA

On the premise that the conductors are nonmagnetic, right-cornered, of such length that end effects are negligible, and that the current density is constant over the cross sections of the individual conductors, the writer[1] has derived an expression, Eq. (1) below, for the inductance of a single phase circuit with conductor cross sections disposed as in Fig. 1. The first and third conditions are commonly satisfied in practice; although rectangular tubular conductors are drawn with rounded edges, to minimize "edge effect," the assumption of right-corneredness[2] is essential to tractable analysis; with practical bus bar spacings the reduction in inductance produced by power-frequency skin and proximity effects is negligible.

$$L=2W/w^2(rs-RS)^2 \text{ abhenries per centimeter of line length;} \quad (1)$$

wherein $W$, the energy of the associated magnetic field, is given by

$$W=w^2[W(r,s)+W(R,S)+2W(t,S+t)+2W(R+t,t)-2W(R+t,S+t)-2W(t,t)]. \quad (2)$$

Typically, $W(r,s)$ is defined by

$$W(r,s)=(\tfrac{1}{6})[2F(D,s)-F(D+r,s)-F(D-r,s)+2F(r,s)-2F(0,s)+4\pi r^2 s(3D-r)]; \quad (3)$$

and, typically, $F(r,s)$ is defined by

$$F(r,s)=(r^4-6r^2s^2+s^4)\log(r^2+s^2)^{\frac{1}{2}}+4rs(r^2-s^2)\tan^{-1}(r/s)-r^4\log|r|. \quad (4)$$

All units are in the absolute system: the current density $w$ in abamperes per square centimeter; $D$, $R$, $S$, $r$, and $s$ in centimeters; $W$ in ergs per centimeter of line length.

In reference 1 are to be found equivalents of (3) especially suited to numerical computation. Since these were derived the writer has discovered (in a paper on transformer design by G. Stein[3]) an extensive table of the function $H(x)$ defined by

$$H(x)=(1-6x^2+x^4)\log(1+x^2)+x^4\log(1/x^2)-8x(1-x^2)\tan^{-1}x. \quad (5)$$

It is not difficult to express $F(r,s)$, and thence $W(r,s)$, in terms of this $H$ function. The resulting expression for $W(r,s)$, evaluated by aid of Stein's table (Table I of this paper), affords the desired check on a value of $W(r,s)$ calculated from a formula of reference 1.

If $r,s\geqslant 0$, as $\tan^{-1}(x)+\tan^{-1}(1/x)=\pi/2$, simple algebraic and trigonometric manipulation of (4) yields

$$F(r,s)=(r^4/2)H(s/r)+s^4\log s-6r^2s^2\log r+2\pi rs(r^2-s^2). \quad (6)$$

[1] T. J. Higgins, Trans. A. I. E. E. **60**, 1046 (1941).

[2] Semi-empirical correction factors are available for approximating the actual effect of the rounded edges: H. B. Dwight and T. K. Wang, Trans. A. I. E. E. **57**, 762 (1938).

[3] G. Stein, Zeits. f. angew. Math. u. Mechanik **9**, 23 (1929).

Similarly, it is easily established that

$$F(r, s) = (s^4/2)H(r/s) + (s^4 - 6r^2s^2) \log s. \quad (7)$$

If $r > s$, (6) is used; if $r < s$, (7) is used; if $r = s$, (6) and (7) are identical.

Substituting (6) in (3) and collecting terms yields

$$\begin{aligned} W(r, s) = (\tfrac{1}{12})[&2D^4H(s/D) \\ &-(D+r)^4H(s/\{D+r\}) \\ &-(D-r)^4H(s/\{D-r\}) \\ &+2r^4H(s/r)] - s^2[2D^2 \log D \\ &-(D+r)^2 \log (D+r) \\ &-(D-r)^2 \log (D-r) \\ &+2r^2 \log r] - 2\pi rs^3/3. \end{aligned} \quad (8)$$

Again, substituting (7) in (3) and collecting terms yields

$$\begin{aligned} W(r, s) = (s^4/12)[&2H(D/s) \\ &-H(\{D+r\}/s) - H(\{D-r\}/s) \\ &+2H(r/s)] + 2\pi r^2 s(3D - r)/3. \end{aligned} \quad (9)$$

### 2. AN ILLUSTRATION

The following example, the data of which are taken from reference 1, illustrates application of (8) or (9):

*Given:* that $r = s = 2.5$ inches, $t = 0.5$ inch, $D = 10$ inches; whence by (2)

$$\begin{aligned} W/w^2 = &W(2.5, 2.5) + W(1.5, 1.5) \\ &+2W(0.5, 2) + 2W(2, 0.5) \\ &-2W(2, 2) - 2W(0.5, 0.5) \\ = &171.203 + 27.354 + 11.547 + 11.522 \\ &-154.528 - 0.950 \\ = &66.153. \end{aligned} \quad (10)$$

*To check:* the value, +154.528, given for the term $2W(2, 2)$, this term having been calculated by (37) of reference 1.

*Check:* By (8)

$$\begin{aligned} W(2, 2) = (\tfrac{1}{12})[&20{,}000H(0.2) - 20{,}736H(\tfrac{1}{6}) \\ &-4{,}096H(0.25) + 32H(1)] \\ &-4[200 \log 10 - 144 \log 12 \\ &-64 \log 8 + 8 \log 2] - 32\pi/3. \end{aligned} \quad (11)$$

As values of the same order and of considerable magnitude are to be added algebraically in (11) values of the $H$'s and of the natural logarithms must be correct to at least five decimal places. Accordingly, Table I cannot be interpolated for the value of $H(\frac{1}{6})$; it must be computed from (5). Evaluated from its power series, $\tan^{-1}(\frac{1}{6}) = 0.165149$ radian. Substituting in (11) the values

| | | | |
|---|---|---|---|
| $H(0.2)$ | $= -0.26818$ | $\log 10$ | $= 2.3025851$ |
| $H(\frac{1}{6})$ | $= -0.18846$ | $\log 12$ | $= 2.4849066$ |
| $H(0.25)$ | $= -0.41038$ | $\log 8$ | $= 2.0794415$ |
| $H(1)$ | $= -2.77259$ | $\log 2$ | $= 0.6931472$ |

and performing the indicated computations on a calculating machine, we have

TABLE I.

| $x$ | $H(x)$ | $\Delta H(x)$ | $x$ | $H(x)$ | $\Delta H(x)$ |
|---|---|---|---|---|---|
| 0.00 | −0.00000 | −0.00070 | 0.51 | −1.44657 | −0.04467 |
| 0.01 | −0.00070 | −0.00210 | 0.52 | −1.49124 | −0.04467 |
| 0.02 | −0.00280 | −0.00349 | 0.53 | −1.53591 | −0.04460 |
| 0.03 | −0.00629 | −0.00488 | 0.54 | −1.58051 | −0.04451 |
| 0.04 | −0.01117 | −0.00626 | 0.55 | −1.62502 | −0.04437 |
| 0.05 | −0.01744 | −0.00764 | 0.56 | −1.66939 | −0.04419 |
| 0.06 | −0.02507 | −0.00900 | 0.57 | −1.71358 | −0.04396 |
| 0.07 | −0.03407 | −0.01035 | 0.58 | −1.75754 | −0.04369 |
| 0.08 | −0.04442 | −0.01169 | 0.59 | −1.80123 | −0.04337 |
| 0.09 | −0.05611 | −0.01301 | 0.60 | −1.84460 | −0.04301 |
| 0.10 | −0.06912 | −0.01432 | 0.61 | −1.88761 | −0.04260 |
| 0.11 | −0.08344 | −0.01561 | 0.62 | −1.93020 | −0.04215 |
| 0.12 | −0.09906 | −0.01689 | 0.63 | −1.97236 | −0.04165 |
| 0.13 | −0.11595 | −0.01814 | 0.64 | −2.01400 | −0.04111 |
| 0.14 | −0.13409 | −0.01938 | 0.65 | −2.05511 | −0.04052 |
| 0.15 | −0.15347 | −0.02060 | 0.66 | −2.09563 | −0.03987 |
| 0.16 | −0.17407 | −0.02179 | 0.67 | −2.13550 | −0.03920 |
| 0.17 | −0.19586 | −0.02297 | 0.68 | −2.17470 | −0.03846 |
| 0.18 | −0.21882 | −0.02411 | 0.69 | −2.21315 | −0.03768 |
| 0.19 | −0.24294 | −0.02524 | 0.70 | −2.25084 | −0.03686 |
| 0.20 | −0.26818 | −0.02634 | 0.71 | −2.28770 | −0.03597 |
| 0.21 | −0.29452 | −0.02742 | 0.72 | −2.32367 | −0.03505 |
| 0.22 | −0.32194 | −0.02847 | 0.73 | −2.35872 | −0.03407 |
| 0.23 | −0.35040 | −0.02949 | 0.74 | −2.39279 | −0.03308 |
| 0.24 | −0.37989 | −0.03049 | 0.75 | −2.42587 | −0.03199 |
| 0.25 | −0.41038 | −0.03145 | 0.76 | −2.45786 | −0.03087 |
| 0.26 | −0.44183 | −0.03239 | 0.77 | −2.48873 | −0.02969 |
| 0.27 | −0.47421 | −0.03330 | 0.78 | −2.51841 | −0.02849 |
| 0.28 | −0.50751 | −0.03417 | 0.79 | −2.54691 | −0.02721 |
| 0.29 | −0.54168 | −0.03502 | 0.80 | −2.57411 | −0.02588 |
| 0.30 | −0.57670 | −0.03584 | 0.81 | −2.59999 | −0.02454 |
| 0.31 | −0.61253 | −0.03661 | 0.82 | −2.62452 | −0.02310 |
| 0.32 | −0.64914 | −0.03736 | 0.83 | −2.64762 | −0.02164 |
| 0.33 | −0.68651 | −0.03808 | 0.84 | −2.66926 | −0.02008 |
| 0.34 | −0.72458 | −0.03876 | 0.85 | −2.68934 | −0.01853 |
| 0.35 | −0.76335 | −0.03941 | 0.86 | −2.70787 | −0.01689 |
| 0.36 | −0.80275 | −0.04002 | 0.87 | −2.72476 | −0.01524 |
| 0.37 | −0.84278 | −0.04060 | 0.88 | −2.74001 | −0.01347 |
| 0.38 | −0.88337 | −0.04114 | 0.89 | −2.75347 | −0.01171 |
| 0.39 | −0.92451 | −0.04164 | 0.90 | −2.76518 | −0.00988 |
| 0.40 | −0.96615 | −0.04211 | 0.91 | −2.77506 | −0.00798 |
| 0.41 | −1.00826 | −0.04254 | 0.92 | −2.78304 | −0.00606 |
| 0.42 | −1.05080 | −0.04293 | 0.93 | −2.78910 | −0.00407 |
| 0.43 | −1.09373 | −0.04329 | 0.94 | −2.79317 | −0.00210 |
| 0.44 | −1.13701 | −0.04360 | 0.95 | −2.79518 | +0.00015 |
| 0.45 | −1.18061 | −0.04388 | 0.96 | −2.79503 | +0.00221 |
| 0.46 | −1.22448 | −0.04411 | 0.97 | −2.79281 | +0.00444 |
| 0.47 | −1.26859 | −0.04430 | 0.98 | −2.78838 | +0.00675 |
| 0.48 | −1.31289 | −0.04445 | 0.99 | −2.78162 | +0.00904 |
| 0.49 | −1.35734 | −0.04459 | 1.00 | −2.77259 | |
| 0.50 | −1.40193 | −0.04464 | | | |

$$W(2,2)=(\tfrac{1}{12})[-5{,}363.60+3{,}907.91+1{,}680.92-88.72]-4[460.5170-357.8266-133.0843+5.5451]-33.5103=11.375+99.3945-33.5103=77.26. \quad (12)$$

This value 77.26 agrees with $154.528/2=77.264$ to 1 in 19,000. It is as exact a check as Stein's five-place table, when utilized as above, allows. Calculation in the fashion

$$\{[20{,}000H(0.2)-20{,}736H(\tfrac{1}{6})]=20{,}000[H(0.2)-H(\tfrac{1}{6})]-736H(\tfrac{1}{6})\}$$

yields greater accuracy; but such refined calculation is not commonly necessary.

## Fatigue of Fabrics

W. F. Busse, E. T. Lessig, D. L. Loughborough, and L. Larrick
*The B. F. Goodrich Company, Physical Research Laboratory, Akron, Ohio*

(Received July 13, 1942)

The effect of temperature and time of application of load on the growth and the apparent "tensile strength" of fabrics was studied with a number of new tests. The results gave considerable information about the mechanism of failure of mechanical fabrics at low loads in service which could not be obtained with standard textile tests.

A "fatigue" test was developed which measured the lives of samples at elevated temperatures when subjected to a constant average load and a superimposed cyclic stress. The change in life with temperature on this test is enormously greater than the change of tensile strength with temperature. Over a considerable range the logarithm of the life under a given load is a linear function of the reciprocal of the absolute temperature, showing that the failure depends on a viscous or plastic flow within the fibers. Calculated activation energies for this flow process increase for different fibers in the order rayon, cotton, Nylon. The construction of the fabric affects the absolute life, but not the calculated activation energy. When creep is negligible, a loaded cord contracts when the temperature is raised, analogous to the Joule effect in rubber.

### I. INTRODUCTION

THE service life of mechanical fabrics such as belt ducks, tire cords, and sewing threads, as well as the wear resistance of apparel and seat-cover fabrics, commonly is limited by the ability of the fabric to withstand repeated small stresses which are far below the nominal "tensile strength" of the material. The ability to withstand these repeated small stresses may be designated by the broad term, "fatigue resistance."

For a long time the only method, and even today the most common laboratory method, of evaluating the serviceability of mechanical fabrics is to measure the "tensile strength." This property is not only used as a guide to the value of the final product, but it is also used both in scientific and industrial laboratories as an important measure of the quality of the cotton used in making the product.[1]

A great deal of work has been done standardizing tensile tests[2] since it was early recognized that tensile strength measurements of fabrics change not only with the humidity and temperature, but also with other variables such as sample dimensions, rate of loading, etc. The various committees working on this problem not only succeeded in standardizing the testing equipment and methods so that different laboratories could check each other's results, but they did it so well that there is today some tendency to consider tensile tests measured under standard conditions of 72°F and 65 percent rh and specified rate of loading almost as something final and absolute, and to consider measurements under other conditions as unimportant, if not actually improper.

Any standardization of thinking on this subject tends to make us minimize or entirely overlook the fact that the tensile strength measured under standard conditions is not a satisfactory measure of the quality or serviceability or "fatigue resistance" of mechanical fabrics any

[1] Underwood, Shirley Inst. Mem. **14**, 1 (1935); Campbell, Text. Research **8**, 263 (1938); Webb, A.S.T.M. Committee Meeting D-13, 404 (1940).

[2] See, for example, reports of A.S.T.M. Committee Meeting D-13 (1920–1942); Peirce, Urquhardt, and Midgley, Shirley Inst. Mem. **2**, 109–150 (1923).

more than tensile tests of metals are a measure of their fatigue resistance. Also, it makes us overlook the fact that the mechanical properties of the fibers themselves (length, fineness, maturity, bundle strength) do not correlate too well with the standard tensile properties of yarns or cords made from them. In fact, we sometimes even fail to recognize that the "simple" act of breaking a yarn or cord is a very complex process, perhaps even more complex than the process of breaking a wire—a phenomenon which has received a great deal of detailed study by the metallurgists. If a question is raised about the mechanism of the failure of the yarn or cord under tensile stress, it usually is answered (with finality) by the statement that it was due partly to slippage and partly to breakage of the fibers. This too often closes the inquiry, even though there is no agreement in the literature as to the relative amounts of slippage and breakage of fibers which occur under various conditions and there is almost no information about the mechanism of the breakage of the fibers themselves.

In the present work measurements of tensile strength were made over a wide range of rates of loading, temperature, etc., to study some of the factors which are masked by present standard tensile tests and which are important in determining the so-called "fatigue resistance" of cords in actual service.

As a result of this work a new fatigue test was developed which gives some additional information about one type of "fatigue life" of cord structures and also about the fundamental mechanisms causing the failure of fabrics under tensile stresses.

## II. PREVIOUS WORK

The effect of rate of loading on the tensile strength of yarns and fibers was studied by Midgley and Peirce[3] in excellent pioneering research. They showed that the apparent "tensile strength" of a yarn changed about 10 percent of its nominal value for each tenfold change in the rate of loading, over the range of loading times from 0.02 second (on the ballistic pendulum) to about 18 days. A yarn that had a "tensile strength" of 500 grams on the ballistic test would fail at a load of 200 grams if the load was applied over four days. However, the ultimate elongation was reported to be about the same in all tests. The strength of single cotton fibers varied with the rate of load by the same factor as did the yarns. Since the yarns had only about half the total tensile strength of the fibers in their cross section, slippage of fibers did not play an appreciable part in this change of strength of yarn with rate of loading.

Lutts and Himmelfarb[4] found that ropes would fail at loads far below their "tensile strength" if the loads were applied for long times. They observed that the change in life with load was different for cotton and Manila ropes.

Castricum and Benson,[5] who studied the effect of rate of loading on rayon yarns, found the "tensile strength" to vary with logarithm of the rate of loading in agreement with Midgley and Peirce. This logarithmic relation was also consistent with the results of Bellinson[6] who, however, expressed his results in a different way. Since they worked with continuous filament yarns, their results must have been due to the flow within the fibers and not to slippage. Leaderman[7] has started a fundamental study of the growth, recovery, and elastic after-effects of fibers submitted to intermittent loads, but he says nothing about the effect of rate of loading on "tensile strength."

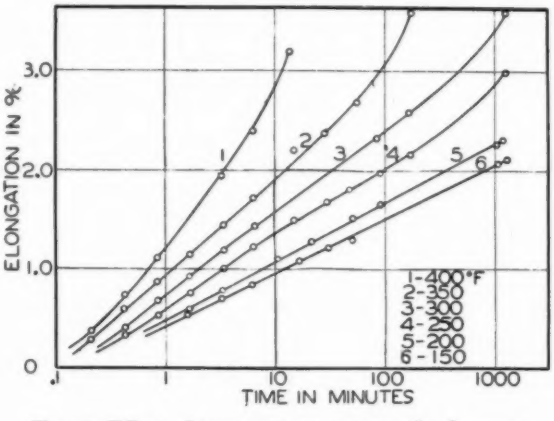



FIG. 1. Effect of temperature on growth of rayon.

None of these authors suggests that there is any difference in the kind of failures at low and high rates of loading, or that the change in

[3] Midgley and Peirce, Shirley Inst. Mem. **5**, 102 (1926).

[4] Lutts and Himmelfarb, Proc. A.S.T.M. **40**, 1251 (1940).

[5] Castricum and Benson, Proc. A.S.T.M. **41**, 1214 (1941).

[6] Bellinson, Text. Research **10**, 316 (1940).

[7] Leaderman, Text. Research **11**, 171 (1941).

strength with rate of loading might be some measure of a fundamental property of the material.

Few, if any, data are available in the literature on the effect of a wide range of temperatures and humidities on tensile strength. There are some data on the amount of water taken up from atmospheres of different humidities at different temperatures. "Bone dry" tensile tests often are made, but their significance is open to some question.

The effect of aging at high temperatures has been studied but tensile strengths at these high temperatures usually are not measured.[8]

### III. EXPERIMENTAL WORK

Since the literature contains little or nothing on the effect of a wide range of temperatures on the tensile and elastic properties of fabrics, a number of tests were made to study some of these relations.

#### 1. Growth of Cords at Various Temperatures and Loads

The change in length of cords with time at various loads and temperatures was measured in the hope that this would show something of the mechanism of failures. Tests were made by passing the cord through an electrically heated oven, loading the cord with a dead weight (subject only to building vibration), and measuring the change in length with a dial gauge. Typical results for a 275 denier/4/3 high tenacity rayon cord at various temperatures are shown in Fig. 1. The cord had about a 20-lb. tensile strength at room temperature and was loaded with a 4-lb. weight. At any one temperature the elongation is about proportional to the logarithm of the time over the range from 0.1 to 1000 minutes, provided the total elongation is less than 2.5 percent. The rate of growth increases slightly as the temperature is raised.

If the temperature is kept constant but the load is varied, a similar set of curves is obtained, as shown in Fig. 2. While these curves do not have any high precision due to the difficulty of

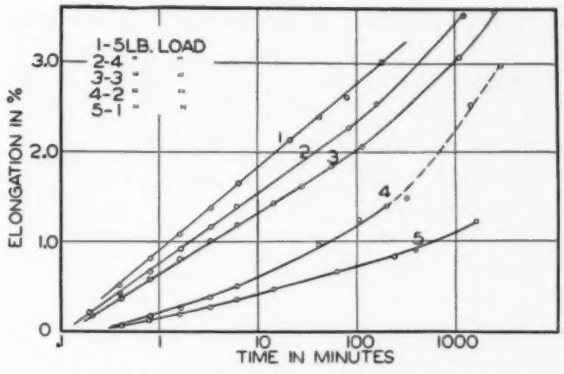



FIG. 2. Growth of rayon cords at 300°F.

determining the zero point, the results show the trends with sufficient accuracy. Similar curves can be obtained with cotton cords, but the greater variability of the cotton makes it more difficult to obtain reproducible results.

The growth of the rayon cord cannot be due to slippage of the continuous filaments past each other, after the initial consolidation of the cord, so it must be due to a stretching, or plastic flow, of the fibers. Raising the temperature then reduces the viscosity (or plasticity) of both rayon and cotton fibers. This effect later turns out to be of considerable importance in relation to the "fatigue" properties of these cords.

Another interesting effect of temperature on the elongation of cords was observed in this work. After the cords had been under load for some hours, so the rate of growth was small, an increase in temperature produced an immediate and reversible *decrease* in length. Figure 3 shows the results when a 22's/5/3 cotton cord with a nominal tensile strength of 19 lb. at room temperature was held under 0.7-lb. load for two hours at 300°F, and the temperature was then varied from 200 to 300°F several times. The initial temperature increase sometimes did not show this decrease in length, due, perhaps, to changes in moisture. This factor may be responsible for the fact that Meyer and Lotmar did not find this type of temperature effect.[9] Figure 4 shows curves for a 162 denier/5/3-SSZ Nylon cord loaded to 7 lb. Here the magnitude of the effect is much greater. These data were taken by Messrs. B. D. Morgan and M. W. Wilson of this laboratory.

The contraction of the cotton cord is about

[8] See U. S. Pat. 2,050,196 and 2,050,197 (Sebrell) for data on effect of various chemicals on tensile strength at 75° and 250°F. Also Valko (Milliland's Textilberichte No. 9, 1932) measured properties of cotton down to liquid-air temperatures (reported by Meyer and Lotmar, Helv. Chim. Acta **19**, 68 (1936)).

[9] Meyer and Lotmar, Helv. Chim. Acta **19**, 68 (1936).

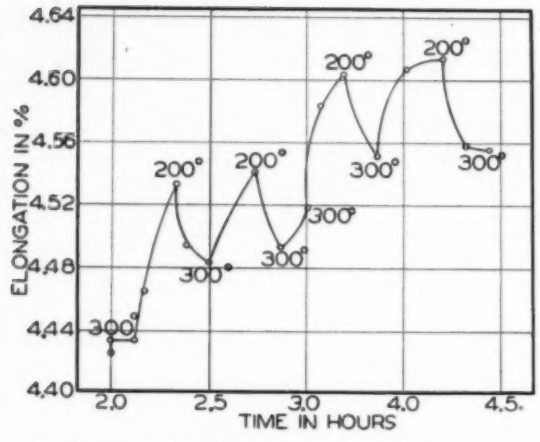



FIG. 3. Change in length of cotton cord with temperature.

0.06 percent for a 100°F increase in temperature and the corresponding value for Nylon is about 0.6 percent change in length per 100°F temperature change. These changes are opposite in direction to those produced by simple thermal expansion. They might conceivably be produced by an anisotropic expansion of the fibers, just as a cord contracts when wet due to the lateral expansion of the fibers, but it is more probable that these changes are due to an increase in the modulus of the fibers at higher temperatures. This effect is similar in kind, but smaller in amount, to that found in rubber and other high polymers, and can be explained by the kinetic theory of elasticity.[10]

## 2. Effect of Temperature on Tensile Strength

Tensile tests were made at elevated temperatures on a 275 denier/4/3 rayon tire cord, and on a 22's/5/3 cotton tire cord and a commercial wet twisted cotton cord.

A small cylindrical oven was built to go between the jaws of a tensile machine, so the cords could be broken at the high temperatures. The air in the oven had the absolute moisture content to give 65 percent rh at 72°F. The cords were held at the high temperature for three minutes before testing, which was long enough to reach temperature equilibrium, but not long enough to get appreciable deterioration at the higher temperatures.

The results shown in Fig. 5 indicate the combined effects of temperature and drying. The values for rayon at 150° and 200°F are uncertain because of a tendency of the cords to break in the grips at those temperatures and there is considerable statistical variation in the results for the cotton cords. At 200°F these particular rayon and cotton cords lose about 7 percent and 20 percent, respectively, of their normal tensile strengths. At 350°F the corresponding losses are about 20 percent and 35 percent for rayon and cotton.

[10] See Meyer, *Physical Chemistry of High Polymeric Systems* (New York, 1940) for a review of this subject.

## 3. Life of Cords under Load

### (*A*) *With Small Vibrations*

Midgley and Peirce studied the effect of rate of loading on tensile strength at room temperature. In the present work the life of cords was measured under various loads and at elevated temperatures. For preliminary work an oven was made 20″ long, 5″ high, and ¾″ wide, through which 16 cords could be hung from grips above the oven. Loads were then hung from the lower end of the cords. The upper grips were supported from a bar fastened to a heavy steel table which rested on a floor which was subject to some vibration from factory equipment running in the same building.

The effect of load on the life at 280°F of 8.5″ lengths of two cords of 1¼″ staple cotton and 22's/5/3 construction, but of different twists, is shown in Fig. 6. Both cords have about 20-lb. "tensile strength."

Each point is the average of only 8 cords tested, so there is some scattering of the data. However, it is seen that the life is increased by a factor of about 10 for each 2-lb. decrease in load. The log of the life is a linear function of

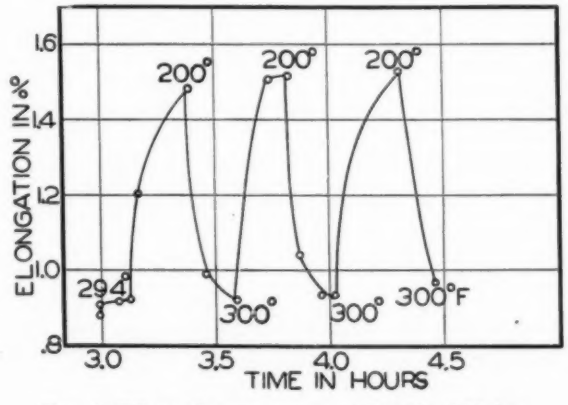



FIG. 4. Effect of temperature on length of Nylon.

the load over the range from about 15 to 5 lb. or for lives of a few seconds up to about a week. This is the same form of curve as that found by Midgley and Peirce. The slopes of the curves are the same for the two cords of different twists, although the absolute values of the lives at any one load differ by about two to one.

The equations of the curves for the two cords are:

$$W = -1.6 \log L + 11.4$$

and

$$W = -1.6 \log L + 12.1,$$

where $W$ = load in pounds, $L$ = life in minutes.

It might be thought that the failure of cords under low loads in long times was due to progressive chemical degradation of the fibers at the high temperatures. However, other data shown below throw doubt on this conclusion, and it does not explain similar data obtained at room temperatures.

The linear relation between log life and load cannot be valid to extremely long times since this would require the cord to fail at zero load in a finite time (about 100 years). Midgley and Peirce's yarn would fail at zero load in about $10^{11.1}$ seconds or 4000 years.

It seems clear from these data that the mechanism of failure at lives around 1 to 100 hours changes smoothly to the mechanism of failure in the standard tensile tests as the rate of loading is increased. Writing the above equation in the form $W = A \log L + B$, $A$ might be considered a viscosity factor and $B$ a structure factor, which may, perhaps, indicate the distribution of stresses among the fibers, etc.

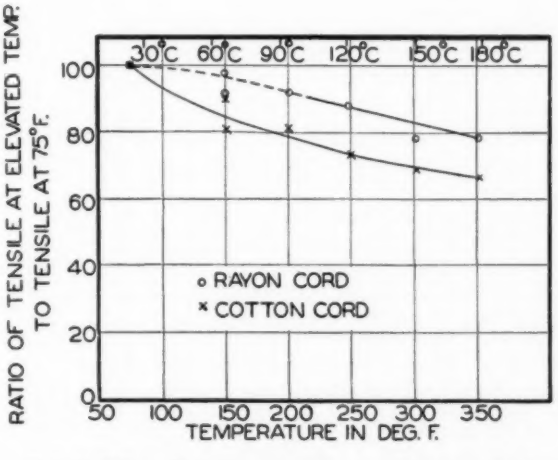



FIG. 5. Effect of temperature on tensile strength of cotton and rayon cords.

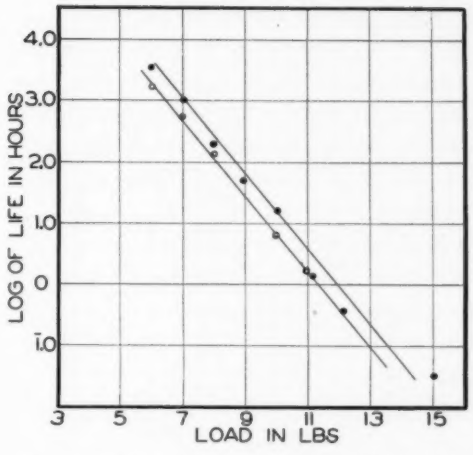



FIG. 6. The effect of dead load on the life of two cotton cords.

TABLE I.

| Length of cord between grips | Life in hours Cord *A* | Cord *B* |
|---|---|---|
| $\frac{1}{4}''$ | 37 | 52 |
| $\frac{1}{2}''$ | 29 | 50 |
| 1″ | 37 | 37 |
| 2″ | 35 | 49 |
| 4″ | — | 49 |
| 6″ | — | 45 |
| $8\frac{1}{2}''$ | 35 | 49 |

Both fiber slippage and fiber breakage have been suggested as causes of cord failure in tensile tests. If they both operate in the dead load tests, then one should find the life of a cord to increase rapidly as the length of the sample between grips was decreased to less than the fiber length. The cords *A* and *B* of Fig. 6 were tested with various lengths between the grips and under 6-lb. load at 280°F. The results given in Table I show the life to remain roughly constant on going from a sample length of $\frac{1}{4}''$ to $8\frac{1}{2}''$. The failure of the cords must then depend primarily on some factor other than slippage of the fibers.

### (*B*) *With Cyclic Stresses*

(1) *Description of test.*—In service most fabrics are subject to stresses which change over a rather wide cycle, instead of to the nearly constant stresses such as were used above. It is known that the range of the stress cycle plays an important part in the fatigue life of a sample, just

as does the average stress, so the machine (U. S. Patent 2,240,505) shown in Fig. 7 was built to test cords under cyclic stresses.[11] This test has the additional advantage over the "dead load" test that it is much less sensitive to the occasional shocks and large vibrations which may occur in any building.

Cords about 8" long are hung from the upper bar, which is vibrated, usually with a stroke of $\frac{1}{32}$" or $\frac{1}{16}$", by means of eccentric bearings. The cords pass through a vertical, electrically heated oven and weights are hung from the lower ends of the cords. The oven is like that described above for dead load test, and is controlled by a potentiometer controller.

TABLE II.

| | Life of cords in hours | |
|---|---|---|
| | Cotton | Rayon |
| First break | 0.5 | 11.0 |
| 10% percentile | 3.4 | 16.2 |
| 25% percentile | 5.2 | 17.8 |
| 50% percentile | 8.0 | 19.2 |
| 75% percentile | 11.4 | 21.7 |
| 90% percentile | 16.0 | 24.1 |
| Last cord | 36.6 | 30.0 |
| Average | 8.8 | 18.5 |
| Standard deviation | 3.8 | 3.5 |

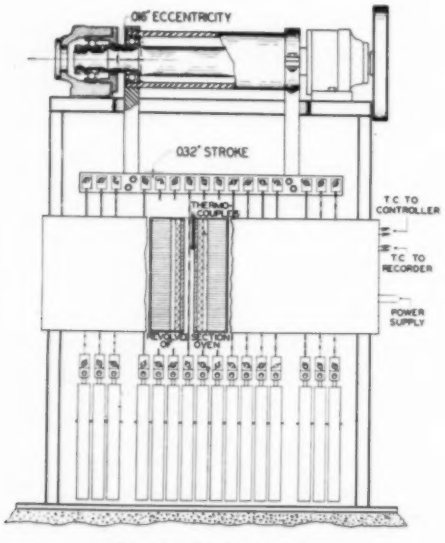



FIG. 7. Cord vibrator.

[11] For other more complicated types of "fatigue tests" for mechanical fabrics see Allen and Te Grotenhuis—U. S. Pat. 2,157,092—1936; Castricum—U. S. Pat. 1,923,296—1933; Buffum—U. S. Pat. 1,785,690—1932; Oxley—U. S. Pat. 1,499,546—1924; Candee—U. S. Pat. 1,435,208—1922.

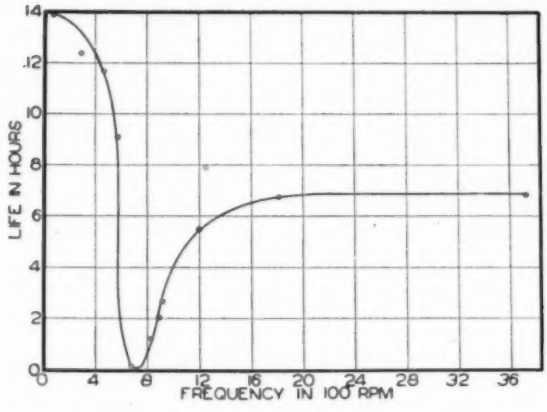



FIG. 8. Effect of impressed frequency on vibration life of a rayon cord.

Thermocouple tests showed some vertical temperature gradient in the oven, but the temperature was about constant for the central 2", and the temperature 1" from top was the same at different places in the oven within about ±3°F. Where this difference was significant, tests were run by hanging control and experimental cords on alternate grips and then repeating the test with the samples reversed. Because the grips were outside the oven, there was very little trouble with grip breaks.

(2) *Effect of frequency.*—It is necessary to keep the frequency of vibration of the upper bar several times greater than the natural frequency of the weighted cords to avoid resonance effects and very large forces on the cords. Figure 8 shows the life of rayon cords (275 denier/4/3 construction, 7½" long, 350°F, 5-lb. load, 0.032" stroke) as a function of the vibration frequency of the upper bar. As long as this frequency is over 2.5 times the resonance frequency the life is about constant. This condition was maintained with most cotton and rayon cords tested, but it did not exist with low stretch glass or wire cords so that they failed in a few seconds on these tests.

(3) *Spread of results.*—As with nearly all fatigue or life tests, the spread of values is quite large. Table II gives the results of a statistical study of the lives of cotton and rayon cords. Five hundred cotton cords of 22's/5/3 construction were tested on three vibrators (5-lb. load at 300°F, $\frac{1}{32}$" stroke and 7½" length). One hundred rayon cords of 275 denier/4/3 construction were

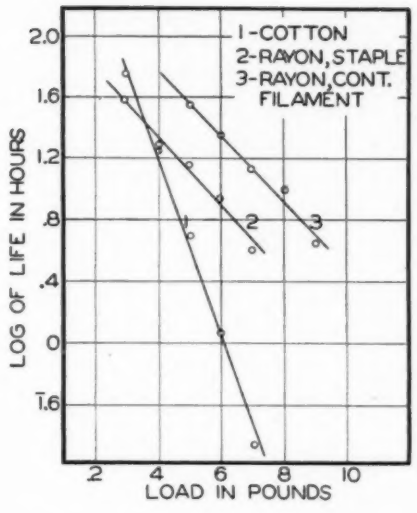



FIG. 9. Effect of load on vibration life.

tested at the same conditions except that the temperature was raised to 320° to make the samples break in a shorter time. The rayon cords are considerably more uniform than cotton cords, the coefficients of variation being about 19 percent for rayon and 43 percent for the cotton cords. The absolute lives vary somewhat with the particular vibrator used, due to slight differences in bearings, etc.

In view of the large spread, and the wide changes in life that can be made by changing the temperature (see below), standard deviations and skewness usually are not calculated. The results are reported usually as the average life of 1 to 4 sets of 8 cords each. In cases where the median differs considerably from the average both values are considered. It is interesting to note that the spread of these results is no greater than the spread of fatigue life tests on precision mechanical items such as notched rods or spiral bevel gears where it is found that there may be over a tenfold range in the fatigue lives at any given load.[12]

(4) *Effect of load on vibration life.*—If the temperature and stroke are kept constant, then the logarithm of the life is found to vary linearly with the load, just as with the dead load test. Figure 9 shows curves for tire cords of cotton, continuous filament high tenacity rayon, and a staple rayon taken under various conditions.

It is significant that the latter two cords have about the same slope, while the cotton cord has a much higher slope. Because of this difference in slopes of cotton and rayon cords, their relative value, as determined by the fatigue life on this test, can be varied over wide limits, depending on testing conditions.

(5) *Effect of stroke and sample length.*—If the dead weight load on the cord is kept constant, and the stroke or range of the vibration cycle is increased, the maximum dynamic load and the stress cycle will be increased. If the sample length is increased, the dynamic force is decreased. Over a certain range, the log of the life is again a linear function of the stroke and length but this cannot be true over a wide range because with small enough vibrations the life is independent of the cord length. The probable relation between life, stroke, and cord length for a given load is illustrated more or less schematically in Fig. 10.

Since the dynamic force on the cord is proportional to stroke and inversely proportional to cord length, the life data can be plotted against the reduced stroke, $K$=stroke/length, to get a curve of the form of Fig. 11. One might postulate that the nearly horizontal bottom part of the curve is determined by the tensile properties of the cords, while the more nearly vertical part, where the lives are long, is related to the "fatigue" properties of the cords.

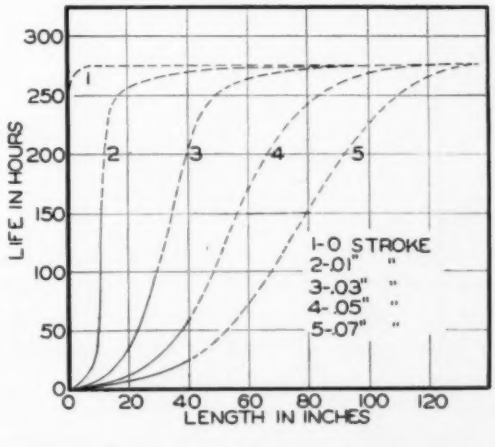



FIG. 10. The effect of stroke and length on life.

[12] Alman, "Facts and fallacies in stress determination," S. A. E. J. p. 52 (1942). See also Macauley, "Endurance of ball bearings," Automotive Eng. **13**, 213–223 (1923); Styri, Mech. Eng. **47**, 490–492 (1925).

TABLE III.

| Initial vibration period (hours) | Vibration life second period (hours) | Total vibration life (hours) |
|---|---|---|
| 0 | 14.7 | 14.7 |
| 5 | 7.4 | 12.4 |
| 10 | 4.1 | 14.1 |
| 15 | 2.0 | 17.0 |

(6) *Mechanism of failure*

(a) Irreversible loss of fatigue resistance during vibration.—Groups of cords were vibrated ($\frac{1}{16}$″ stroke, 13″ length) for periods of 0, 5, 10, and 15 hours, and removed from the vibrator. After a few days each set was put back on the vibrator and run to failure. The second part of the tests was run on samples 12″ long to avoid having to put the same parts of the cord back into the grips. The results were as shown in Table III.

The total lives in these tests are roughly constant. The high total life of the last sample is due to the fact that only the better cords survived the first 15-hour treatment, and so they brought up the total life.

Other tests showed that the cords could be stored for a period of a week to a month without recovering their vibration life. Treatments such as mechanical flexing, or wetting with water, or extracting with solvents did not restore the fatigue life which had been lost.

(b) Heat aging *vs.* vibration life.—These results might be thought due to merely a heat aging of the samples. This cannot be the whole explanation of the results, however, for the loss in tensile strength of the cords is more rapid when the samples are being vibrated than when the samples are merely held in the oven at the high temperature.

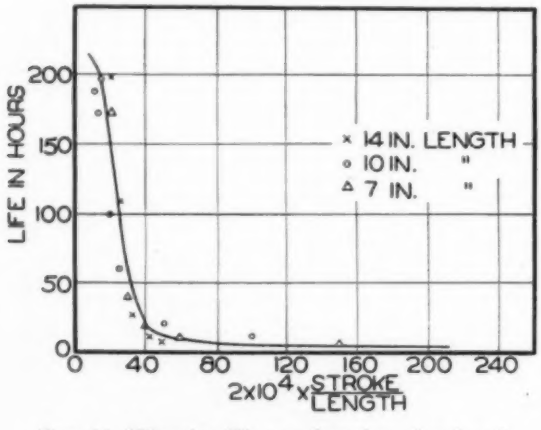



FIG. 11. Vibration life as a function of reduced stroke (stroke/length).

One set of cords (22/5/3–1¼″ staple cotton) was held in the oven at 280°F for 25 hours under a 5-lb. load without any vibration, and the tensile strength determined at room temperature. Other cords were treated similarly except that during part of the time the upper ends of the 13″ long cords were subject to a $\frac{1}{32}$″ vibration at 3600 cycles per minute. The results are shown in Table IV.

This shows conclusively that the deterioration during vibration at high temperatures is not merely a heat aging, but involves some other factors.

(c) Temperature coefficient of vibration life.—Figure 12 shows the changes in tensile strength and vibration life with temperature for a cotton and a rayon cord. It is seen that the temperature coefficient of vibration life is enormously greater than the temperature coefficient of the tensile strength. This is further proof that vibration life depends on some factor other than that controlling the tensile strength.

Figure 13 shows the results of another series of tests on a 22's/5/3 cotton cord at different temperatures (5-lb. load, 0.032″ stroke, 7.5″ length), the logarithm of the life being plotted against the reciprocal of the absolute temperature. Sixteen cords were tested at each temperature. There is some question about how to average the results, so the arithmetic average and the logarithmic average and the median of each test were plotted. It is seen that within the range of these different averages, the logarithm of the life is a linear function of the reciprocal of the absolute temperature for more than a tenfold change in life, or

$$\log L = (A/T) + B,$$

where $L$ = life in hours, $T$ = absolute temperature, and $A$ and $B$ are constants.

This surprising and important relation might suggest that the fatigue life is determined by chemical degradation at high temperatures, since chemical reactions usually have this form of temperature coefficients. However, this cannot be correct since at any one temperature tensile tests merge smoothly into life tests as the applied load is reduced. Neither is this view consistent

with the fact that aging under load at a given temperature is not equivalent to vibrating the cord at that temperature in its permanent effect on the tensile strength of the cord.

If the linear relation between log life and $1/T$ is not due to a chemical reaction, it must be due to some other process involving units which are small enough to be affected by $kT$ ergs of energy. These units obviously cannot be the whole fibers, and surface friction forces do not have this type of temperature coefficient, so fiber slippage cannot be the major factor.

Viscous or plastic flow within the fibers apparently is the only other factor which can account for this relation between vibration life and temperature. If this explanation is correct, Eyring's theory of viscosity[13] can be applied to

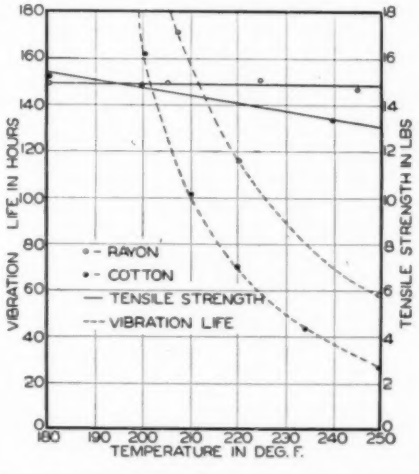



FIG. 12. The effect of temperature on tensile strength and vibration life.

the data of Fig. 13 to calculate for the heat of activation $E$ of the flow process in these cotton fibers, using the equation:

$$\log L = (E/RT) + B$$

where $R$ = gas constant. It is found that $E$ is about 10,300 cal./deg. and $B = -11.3$. Other tests on cotton cords have given values of $E$ ranging from 9,400 to 15,400 cal./deg.

It should be recognized that the difference in the temperature coefficients of tensile strength and of vibration life also implies a change in the slopes of the log life *vs.* load curves with tem-

[13] Glasstone, Laidler, and Eyring, *Theory of Rate Processes* (New York, 1941).

TABLE IV.

| Time with vibration | Time without vibration | Total | Tensile strength | Elongation at 10 lb. |
|---|---|---|---|---|
| 0 hrs. | 0 hrs. | 0 hrs. | 19.0 lb. | 12% |
| 0 | 25 | 25 | 17.66 | 13.41 |
| 2 | 23 | 25 | 15.70 | 12.45 |
| 5 | 20 | 25 | 14.90 | 12.03 |
| 10 | 15 | 25 | 12.34 | 11.99 |
| 20 | 5 | 25 | 12.01 | 11.71 |
| 25 | 0 | 25 | 11.99 | 11.88 |

TABLE V.

| Material | Energy of activation |
|---|---|
| Nylon | 19,870 cal./degree |
| Cotton | 12,520 |
| Rayon | 8,250 |

perature. It also follows that there should be a change in the slopes of the log life *vs.* $1/T$ curves with load, so the calculated "activation energies" must be interpreted with caution.

Eyring's general theory of viscosity states that the rate of viscous flow is proportional to the reciprocal of the relaxation time $\tau$, where

$$\tau \cong 1/\sinh\,(Alf/kT),$$

where $A$ = constant, $l$ = distance moved in 1 jump of the flowing unit, and $f$ = the applied force on the small unit. If $f$ is very large, then

$$\tau \cong e^{-Alf/kT}$$

or at any one temperature

$$\log \tau = C/\log \text{ life} = C_1 \cdot F,$$

where $C$ and $C_1$ are constants, and $F$ = applied force, which is the equation of the life *vs.* load curves of Figs. 6 and 9. It may be mentioned that the rate of growth of some metals also follows the same equation.

At very small values of $f$, the equation extrapolates to the usual equation for viscous flow

$$R = C_2 f,$$

where $R$, the rate of flow, is proportional to $1/\tau$. This explains why the curves of Figs. 6 and 9 cannot be extrapolated to zero load to give a finite life.

Figure 14 gives some preliminary data on the temperature slopes of cotton, rayon, and Nylon. It is seen that the Nylon has the highest tempera-

ture slope, as one might expect from its composition, while cotton is next and rayon is lowest. The values of the calculated heats of activation for the flowing units are given in Table V. Because of the scattering of the data, these values are not too accurate, but they do give some idea of the relative performance of these materials.

A cord of dress-goods rayon had a slope intermediate between that of high tenacity rayon and cotton.

Changes in the construction or twist of cords of any one material changed the absolute vibration life on any particular test, but they had little effect on the slope of the curve obtained by plotting log life against $1/T$. This suggests that the temperature slope under suitable test conditions may be largely a property of the material, while the absolute value of the vibration life is a function both of the material and of the geometry of the cord.

A great deal more work must be done before we get the complete story of what happens in a "simple" operation such as breaking a tire cord in a testing machine, or in the apparently more complex processes of "wearing out" mechanical fabrics in service. The results reported here emphasize one factor that seems to be important in these processes. The many other factors still undiscovered remain as a challenge to stimulate further work in this field.

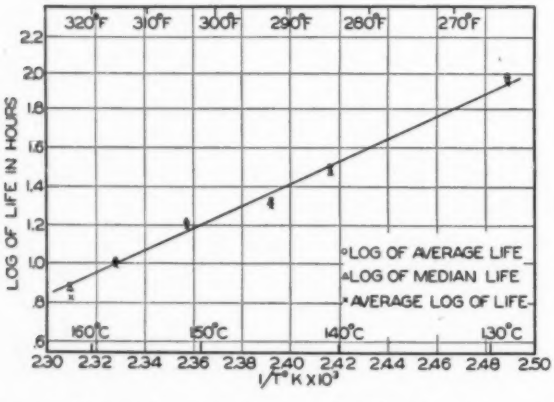



FIG. 13. Effect of temperature on the life of a cotton cord.

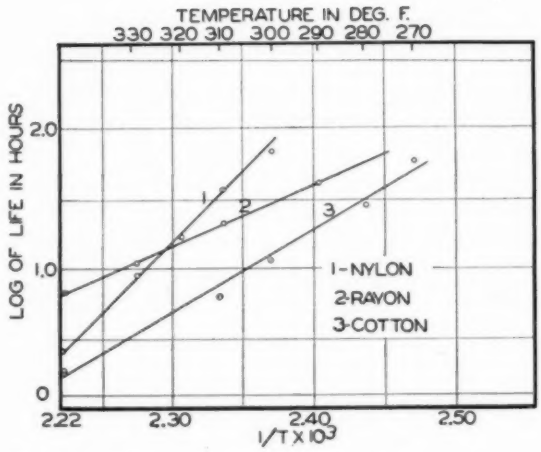



FIG. 14. Change of vibration life with temperature.

## IV. CONCLUSIONS

A study of the behavior of cord fabrics under a wide range of loads and temperatures indicates that the failure of the cords in both ordinary tensile tests and in long time tests is due, in a large part, to some plastic or viscous flow process in the fibers.

The structure of the cord determines the concentration of stresses in the fibers at a given load, and so affects the absolute life under definite testing conditions.

The change of life with temperature is determined largely by the material of which the cord is composed, and is affected relatively little by the cord construction. The temperature coefficient of vibration life may give a measure of the heat of activation for the plastic flow of the fibers.

Under conditions such that flow is negligible, loaded fibers tend to shorten on raising the temperature, similar to the behavior of rubber.

Further work along this line should give some interesting scientific information about the plastic flow of solid fibers. It should also be of considerable practical value in helping the chemists prepare new and better synthetic fabrics to stand up under the long time cyclic stresses and strains which are applied to nearly all mechanical fabrics in service.

# Specific Surface Measurements on Compact Bundles of Parallel Fibers*

R. R. SULLIVAN†
*The University of Tennessee Agricultural Experiment Station, Knoxville, Tennessee*
(Received August 14, 1942)

The flow of air through compact bundles of parallel textile fibers has been studied in order to obtain values of the shape factor $k_0$ for the channels through which the flow takes place. When the fibers are cylinders and the flow is parallel to their axes, $k_0$ is an increasing function of the porosity $\epsilon$. Tables and curves showing this relationship are given. For fibers such as cotton, where the shape and size of cross section changes along the fiber length, $k_0$ changes less rapidly with $\epsilon$. Data are given for this relationship.

## INTRODUCTION

IN a recent review of the permeability method for determining specific surface of fibers and powders,[1] it was pointed out that $k_0$ in the permeability equation took rather small values for parallel flow through compact bundles of parallel circular fibers, and that for the tightest mode of parallel alignment of equal circular cylinders ($\epsilon=0.093$) the value of $k_0$ appeared to be 0.81. This value was contrasted with a fairly constant value of 3.0 to 3.2 for $k_0$ for flow through fibers packed perpendicular to flow but otherwise at random or for particles packed at random over a considerable range of porosities. It is the purpose of the present paper to give data concerning flow through compact bundles of parallel fibers where the flow is parallel to the fiber axes.

The equation for laminar flow of an incompressible fluid through a non-consolidated porous medium has been given[2] as follows:

$$Q=(\zeta A/k_0\mu)(\Delta P/L)(1/S_0^2)[\epsilon^3/(1-\epsilon)^2], \quad (1)$$

where $Q$=volume flow in cm³/sec.; $\zeta\equiv\langle\sin^2\phi\rangle_{Av}$ where $\langle\sin^2\phi\rangle_{Av}$ is the average of $\sin^2\phi$ over the whole of the fluid-solid interface, $\phi$ being the angle a normal to the interface makes with the direction of macroscopic flow; $A$=cross-sectional area of medium in cm²; $\Delta P$=total pressure drop across medium in dynes/cm²; $k_0$=a factor called a shape factor which is the same for all channels having the same geometrical shape; $\mu$=viscosity of fluid in poises; $L$=thickness of medium in cm; $S_0$=specific surface of medium in cm²/cm³; $\epsilon$=porosity of medium=void space divided by total volume. It is $k_0$ of Eq. (1) which is under consideration in this paper.

For the bundles of parallel cylinders to be described here, $\zeta$ takes the value of unity. In the case of bundles of parallel cotton fibers, however, the fiber surfaces, because of convolutions and other irregularities, are oriented in a less simple manner. Thus it is not possible to give a definite value of $\zeta$ for the bundle of cotton fibers. Therefore in this latter case $\zeta$ and $k_0$ will not be separated, but the factor $k$ will be used where

$$k=k_0/\zeta. \quad (2)$$

## MATERIALS STUDIED

### I. Cylindrical Fibers: (a) Goat Wool (Straightened); (b) Chinese Hair; (c) Glass Wool; (d) Blond Hair; (e) Copper Wire; (f) Drill Rods (Two Sizes)

It should be pointed out that in addition to being cylinders these fibers were more or less circular in cross section. The fiber data required for these were values of the specific surface $S_0$ and values of the fiber density $\delta$. In the case of goat wool and blond hair, cross-sectional areas and perimeters were measured with a microscope equipped with camera lucida.* $S_0$ was then given by the ratio

$$\text{sum of fiber perimenters}/\text{sum of cross-sectional areas.} \quad (3)$$

* A portion of this work was presented at the Oxford, Mississippi, meeting of the Southeastern Section of the American Physical Society, April 10–11, 1942.

† Now with Weyerhaeuser Timber Company, Longview, Washington.

[1] R. R. Sullivan and K. L. Hertel, *Advances in Colloid Science* (Interscience Publishers, Inc., New York, 1942), Vol. 1, pp. 37–80.

[2] R. R. Sullivan, J. App. Phys. **12**, 503–508 (1941).

* Thick sections were cut dry in the Hardy sectioning device and viewed in air without embedding fibers in other media.

TABLE I. Fiber data for the cylindrical fibers studied.

| Type of fiber | Specific surface Method | Value of $S_0$ | Density Method | Value of $\delta$ | Equivalent diameter* |
|---|---|---|---|---|---|
| Glass wool | Microscope | 5270 $cm^{-1}$ | Benzene displacement | 2.416 g/cc | 0.00076 cm |
| Goat wool | Microscope | 1028 $cm^{-1}$ | Mineral oil displacement | 1.30 g/cc | 0.00389 cm |
| Blond hair | Microscope | 612 $cm^{-1}$ | Mineral oil displacement | 1.29 g/cc | 0.00654 cm |
| Chinese hair | Perpendicular air-flow | 544 $cm^{-1}$ | Mineral oil displacement | 1.34 g/cc | 0.00735 cm |
| Copper wire | From weight per cm and $\delta$ | 266 $cm^{-1}$ | Hand book | 8.89 g/cc | 0.01505 cm |
| Drill rod | Micrometer caliper | 5.59 $cm^{-1}$ | — | — | 0.716 cm |
| Drill rod | Micrometer caliper | 2.90 $cm^{-1}$ | — | — | 1.379 cm |

* The equivalent diameter was obtained by dividing 4.0 by $S_0$.

For the glass wool, which was quite circular in cross section, diameters were measured with the microscope and $S_0$ was calculated from the relation

$$S_0 = 4\sum d / \sum d^2, \qquad (4)$$

where $\sum d$ = sum of all diameters measured and $\sum d^2$ = sum of squares of these diameters. For the copper wires, the weight per unit length was carefully measured; and, from the known density of drawn copper, the wire diameter and the value of $S_0$ were readily calculated. The value of $S_0$ for Chinese hair was obtained by comparison of perpendicular air-flow measurements with those on other wool which had been measured with a microscope. The value of $S_0$ for the drill rods was calculated from micrometer caliper measurements of their diameters. Densities of goat wool, Chinese hair, and blond hair were obtained by displacement of light mineral oil (sometimes called white oil). In these determinations, which were carried out in an atmosphere of 70°F and 65 percent relative humidity, care was exercised to insure that the oil slowly climbed up into the parallel fiber bundle with no entrapping of air. Sufficient time was given for the weight in the oil to become constant. The density of the glass wool was obtained by displacement of benzene.

Table I gives the fiber data for the cylindrical fibers.

### II. Cotton Fibers: (a) Pima; (b) 190 K $\frac{15}{16}$" Staple; (c) 230 C $1\frac{1}{8}$" Staple

The values of $S_0$ for these cottons were obtained directly by the "arealometer," which is an air-permeability device* using the perpendicular flow principle and having been carefully calibrated over a wide range with cottons which had been measured by the microscope. The values of the density $\delta$ for the cottons were obtained by mineral oil displacement. Table II gives fiber data for the cottons.

* Strictly speaking, the arealometer gives values of $S$, the $cm^2$/gram of fiber, rather than $cm^2/cm^3$. The values of $S_0$ were obtained by multiplying values of $cm^2/g$ by the fiber density. Thus $S_0 = S \cdot \delta$.

### EXPERIMENTAL PROCEDURE

The air-flow measurements were made in a manner similar to that described in a previous paper.[3] Mineral oil flowing from a bottle at constant head caused air to be drawn from the laboratory through the fiber bundle. Measurements were made in a laboratory maintained at 70°F and 65 percent relative humidity. The head was kept so small (from 0.5 to 3.0 cm of kerosene) that correction for the non-compressibility of air was unnecessary. The most loosely compressed bundles of goat wool were contained in a 1.993-cm diameter copper tube of 11.4-cm length without restraining ends, while all other bundles except the drill rods were contained in a 0.831-cm diameter copper tube of 5.05-cm length which also was without restraining ends.

TABLE II. Fiber data for the cotton fibers studied.

| Type of fiber | Specific surface Method | Value of $S_0$ | Density Method | Value of $\delta$ |
|---|---|---|---|---|
| Pima | Arealometer | 5100 $cm^{-1}$ | Mineral oil displacement | 1.56 g/cc |
| 230 C $1\frac{1}{8}$" | Arealometer | 4600 $cm^{-1}$ | Mineral oil displacement | 1.55 g/cc |
| 190 K $\frac{15}{16}$" | Arealometer | 4220 $cm^{-1}$ | Mineral oil displacement | 1.57 g/cc |

In the case of the drill rods flow through only one channel at a time was studied, the channel being formed by binding three or four of the rods tightly together mechanically and then

[3] R. R. Sullivan and K. L. Hertel, J. App. Phys. **11**, 761–765 (1940).

TABLE III. Air-flow data for bundles of parallel fibers where the fibers were cylinders more or less circular in cross section. Fiber data are given in Table I.

| Material | Bundle area $A$ cm² | Bundle length $L$ cm | Volume flow $Q$ cc/sec. | Pressure difference $\Delta P$ dynes/cm² | Bundle mass $M$ grams | Value of porosity $\epsilon$ | Value of factor $k_0$ |
|---|---|---|---|---|---|---|---|
| Goat wool | 3.12 | 11.40 | 207 | 398.7 | 0.721 | 0.9844 | 10.57 |
| Goat wool | 3.12 | 11.40 | 185.4 | 398.7 | 0.772 | 0.9833 | 10.25 |
| Goat wool | 3.12 | 11.40 | 165.8 | 398.7 | 0.832 | 0.9820 | 9.85 |
| Goat wool | 3.12 | 11.40 | 148.0 | 398.7 | 0.902 | 0.9805 | 9.35 |
| Goat wool | 3.12 | 11.40 | 129.6 | 398.7 | 0.983 | 0.9787 | 8.94 |
| Goat wool | 3.12 | 11.40 | 113.1 | 398.7 | 1.082 | 0.9766 | 8.39 |
| Goat wool | 3.12 | 11.40 | 97.7 | 398.7 | 1.202 | 0.9740 | 7.80 |
| Goat wool | 3.12 | 11.40 | 83.9 | 398.7 | 1.350 | 0.9708 | 7.10 |
| Goat wool | 3.12 | 11.40 | 71.0 | 398.7 | 1.544 | 0.9666 | 6.35 |
| Goat wool | 3.12 | 11.40 | 58.6 | 398.7 | 1.803 | 0.9610 | 5.55 |
| Goat wool | 3.12 | 11.40 | 44.8 | 398.7 | 2.162 | 0.9532 | 4.94 |
| Goat wool | 3.12 | 11.40 | 31.5 | 398.7 | 2.705 | 0.9415 | 4.32 |
| Goat wool | 3.12 | 11.40 | 18.79 | 398.7 | 3.602 | 0.9221 | 3.84 |
| Goat wool | 3.12 | 11.40 | 13.97 | 398.7 | 4.323 | 0.9065 | 3.41 |
| Goat wool | 3.12 | 11.40 | 9.40 | 398.7 | 5.405 | 0.8831 | 3.00 |
| Goat wool | 3.12 | 11.40 | 5.70 | 398.7 | 7.213 | 0.8440 | 2.42 |
| Goat wool | 0.542 | 5.10 | 4.52 | 335.7 | 0.353 | 0.902 | 3.06 |
| Goat wool | 0.542 | 5.10 | 4.96 | 440.6 | 0.402 | 0.888 | 2.69 |
| Goat wool | 0.542 | 5.10 | 1.709 | 259.1 | 0.499 | 0.861 | 2.72 |
| Goat wool | 0.542 | 5.10 | 0.853 | 237.3 | 0.702 | 0.805 | 2.06 |
| Goat wool | 0.542 | 5.10 | 0.280 | 380.4 | 1.296 | 0.640 | 1.49 |
| Goat wool | 0.542 | 5.10 | 0.846 | 1194.1 | 1.303 | 0.638 | 1.56 |
| *Blond hair | 0.542 | 5.10 | 2.270 | 869.3 | 1.227 | 0.656 | 1.23 |
| *Blond hair | 0.542 | 5.10 | 0.359 | 830.7 | 1.964 | 0.450 | 1.04 |
| *Chinese hair | 0.542 | 5.10 | 0.346 | 1861. | 2.395 | 0.355 | 1.05 |
| *Chinese hair | 0.542 | 5.10 | 0.347 | 2026. | 2.442 | 0.346 | 0.99 |
| Glass wool | 3.125 | 10.28 | 0.703 | 964.0 | 10.395 | 0.866 | 2.94 |
| *Copper wire | 0.542 | 5.05 | 1.969 | 618.5 | 12.10 | 0.502 | 1.205 |

* Values of $k_0$ for these fibers are corrected for wall effect. (See text.)

sealing the edges with lacquer. A suitable connection was then made for air flow through the channel.

The samples of textile fibers were prepared by tying a thread tightly around one end of a suitably long bundle of parallel fibers and drawing them into the copper tube. A sharp razor blade was then used to cut the fibers even with the ends of the tube. For the more compact bundles of fiber, however, the original parallel bundle was doubled back over a wire at its middle and the wire used to draw the fibers into the copper tube. Suitable bundles of parallel cotton fibers were easily prepared from cotton rovings by means of a "Roll-Spacing Determinator." The Roll-Spacing Determinator is a set of experimental cotton-drafting rolls. Cotton roving is passed through this device with the rolls set for a draft of about ten or twelve and the resulting product allowed to wind on a revolving steel drum. When this ring of fibers is cut at one point and laid out straight a nice bundle of parallel fibers is obtained.

For the coarser fibers, the wall surface of the containing tube was not negligible in comparison to the total fiber surface. Thus in these cases, the tube wall surface was added to the total fiber surface in the tube, the sum being divided by the fiber volume to give a corrected value of $S_0$ to be used in calculation of $k_0$ from the air-flow data.

## RESULTS

From the air-flow data, values of $k_0$ or of $k=k_0/\zeta$ are calculated by means of Eq. (1).

Table III gives air-flow data obtained for the cylindrical type of fiber. Figure 1 shows the relation between $k_0$ and $\epsilon$.

Table IV gives data for the drill rods.

Table V gives air-flow data obtained for the

cotton fibers. In Fig. 2, curve $A$ shows the relation between $k$ and $\epsilon$. The dashed curve $B$ has been added for comparison. Curve $B$ is for the cylindrical fibers.

It is noticed at once that relations for cylinders and for cotton fibers are not the same. One reason for this difference lies in the fact that while $\zeta = 1$ in the first case it is undoubtedly less than unity for the cotton fibers where portions of their surfaces lie at an angle to the direction of macroscopic flow. It also appears likely that the channel shapes are different in the two cases.

## DISCUSSION

In an earlier paper,[2] cases of high porosity bundles were given for which $k_0$, the shape factor for the channels through which the flow took place, became rather large. In the present work, other cases have been given for which values are different from the usual 3.0 to 3.2 found for random packing.[4] It is therefore emphasized that where special types of packing and orientation of particles are employed the value of $k_0$ may be appreciably different from that ordinarily encountered.

From the data presented, it is evident that the air-permeability method of specific surface determination may be applied to compact bundles of parallel fibers, where, however, cognizance must be taken of the fact that $k_0$ is not the same as for some other types of bundles. For certain fibers which are very stiff, or which occur readily in parallel bundles, the parallel-flow type of determination may be simpler than other types.

As an example of specific surface determination by the parallel-flow method, the following data for certain Aralac fibers are given: A mass of 1.185 g of parallel Aralac fibers filling a tube of cross-sectional area 0.542 cm² and length 5.05 cm, constituted the sample. It was found that 185.4 seconds were required for an air-pressure difference of 621 dynes/cm² to cause 50 cc of air to flow through the sample. The density of the

[4] P. C. Carman, J. Soc. Chem. Ind. **57**, 225–234 (1938); **58**, 1–7 (1939).

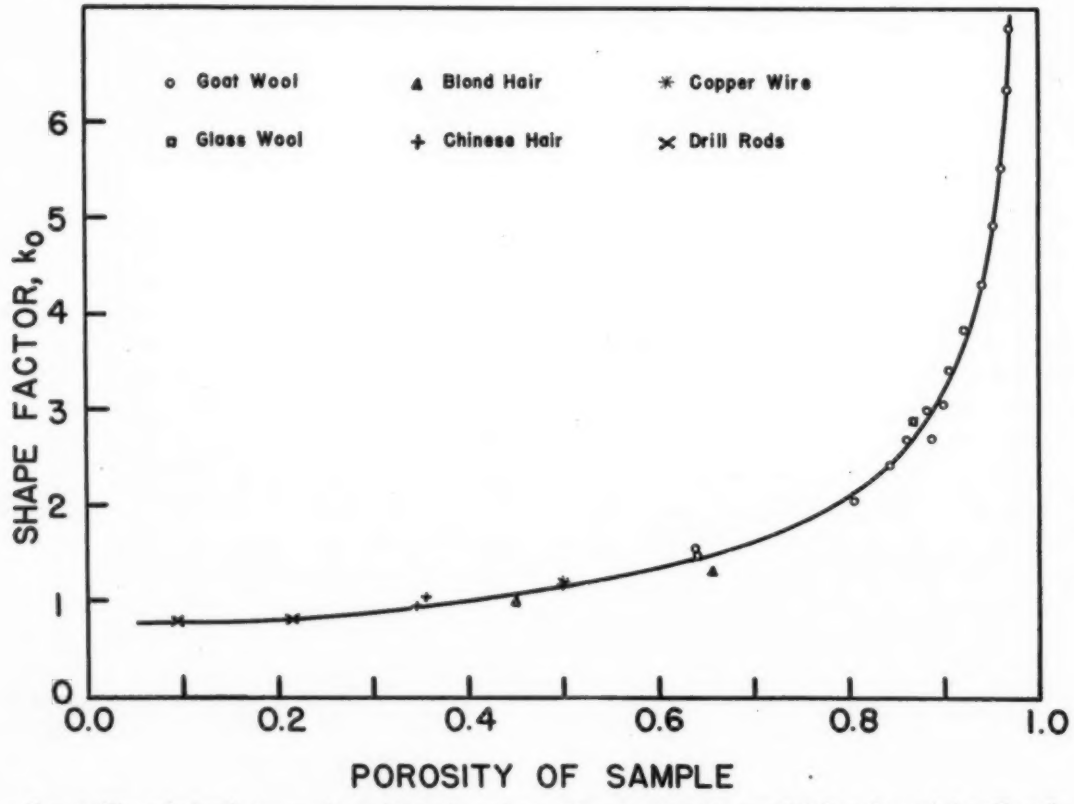



FIG. 1. The relation between shape factor $k_0$ and porosity $\epsilon$ is given for parallel flow through bundles of parallel cylinders which are more or less circular in cross section.

TABLE IV. Data for flow of air through channels formed by uniform parallel drill rods in contact.

| Number of rods used | Rod* radii cm | Rod length $L$ cm | Area† of unit cell $A$ cm² | Area of channel $a$ cm² | Porosity $\epsilon$ $a/A$ | Specific surface $S_0$ cm$^{-1}$ | Volume flow $Q$ cc/sec. | Pressure drop $\Delta P$ dynes/cm² | Value of factor $k_0$ |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.3576 | 92.0 | 0.2215 | 0.02063 | 0.0931 | 5.59 | 0.389 | 766.0 | 0.806 |
| 3 | 0.7145<br>0.6752<br>0.6746 | 91.6 | 0.8196 | 0.0762 | 0.0930 | 2.904 | 5.08 | 745.3 | 0.819 |
| 4 | 0.3577 | 92.0 | 0.5118 | 0.1099 | 0.215 | 5.59 | 4.775 | 254.8 | 0.829 |

* The smaller rods were very nearly equal in size, thus only the average value of their radii is given. The larger rods differed in radii as indicated. The triangle formed by them was not quite equilateral. This was taken into account in the calculation.
† The unit cell for three rods is the triangle formed by their centers. The unit cell for four rods is the square formed by their centers.

TABLE V. Air-flow data for bundles of cotton fibers where the fibers are aligned parallel to the direction of flow.

| Material | Bundle area $A$ cm² | Bundle length $L$ cm | Volume flow $Q$ cc/sec. | Pressure difference $\Delta P$ dynes/cm² | Bundle mass $M$ grams | Value of porosity $\epsilon$ | Value of factor $k$ |
|---|---|---|---|---|---|---|---|
| Pima | 0.542 | 1.98 | 0.206 | 870. | 0.124 | 0.926 | 3.46 |
| Pima | 0.542 | 2.26 | 0.256 | 363.6 | 0.252 | 0.868 | 2.70 |
| Pima | 0.542 | 3.07 | 0.267 | 835. | 0.422 | 0.837 | 2.52 |
| Pima | 0.542 | 2.28 | 0.255 | 1267. | 0.421 | 0.782 | 2.42 |
| Pima | 0.542 | 1.61 | 0.266 | 1303. | 0.337 | 0.752 | 2.35 |
| Pima | 0.542 | 2.29 | 0.1645 | 1261. | 0.495 | 0.744 | 2.34 |
| Pima | 0.542 | 2.29 | 0.1610 | 2089. | 0.588 | 0.696 | 2.32 |
| Pima | 0.542 | 2.27 | 0.0319 | 1898. | 0.872 | 0.545 | 2.31 |
| 230 C | 0.542 | 3.81 | 0.259 | 324.6 | 0.334 | 0.896 | 2.99 |
| 230 C | 0.542 | 5.15 | 0.268 | 722. | 0.565 | 0.869 | 2.75 |
| 230 C | 0.542 | 1.69 | 0.274 | 327.1 | 0.209 | 0.852 | 2.75 |
| 230 C | 0.542 | 3.57 | 0.250 | 699.5 | 0.470 | 0.843 | 2.61 |
| 230 C | 0.542 | 2.93 | 0.262 | 1411. | 0.532 | 0.784 | 2.60 |
| 230 C | 0.542 | 3.10 | 0.202 | 1977. | 0.695 | 0.733 | 2.40 |
| 230 C | 0.542 | 3.10 | 0.118 | 2088. | 0.823 | 0.684 | 2.52 |
| 230 C | 0.542 | 1.65 | 0.104 | 1642. | 0.510 | 0.632 | 2.45 |
| 230 C | 0.542 | 2.23 | 0.0388 | 1760. | 0.833 | 0.555 | 2.41 |
| 190 K | 0.542 | 3.12 | 0.266 | 490.2 | 0.394 | 0.852 | 2.71 |
| 190 K | 0.542 | 2.26 | 0.278 | 691.8 | 0.368 | 0.809 | 2.60 |
| 190 K | 0.542 | 2.27 | 0.292 | 1299. | 0.464 | 0.760 | 2.44 |
| 190 K | 0.542 | 2.29 | 0.0931 | 1811. | 0.728 | 0.626 | 2.43 |
| 190 K | 0.542 | 2.28 | 0.0890 | 1790. | 0.735 | 0.621 | 2.41 |

Aralac fiber was 1.28 g/cc. The porosity $\epsilon$ of the sample was therefore $\epsilon=1-(1.185)/(1.28)(0.542)(5.05)=0.662$. Since Aralac fibers are cylindrical and more or less circular in cross section, the appropriate value of $k_0$ may be obtained from the curve of Fig. 1. This curve gives $k_0=1.50$ for $\epsilon=0.662$. Using $\zeta=1$ for parallel flow and $\mu=185\times10^{-6}$ poise for air, Eq. (1) yields a value of $S_0=1.50\times10^3$ cm$^{-1}$ for the Aralac fibers. In order to check the accuracy of this value a larger quantity of these fibers (2.059 g) was cut in short lengths (about 3 mm) and stacked perpendicular, but otherwise at random, to the axis of a tube of 2.85 cm² cross-sectional area. Perpendicular-flow measurements[1, 5] upon this sample ($\zeta=0.5$, $k_0=3.0$) yielded a value of $S_0=1.47\times10^3$ cm$^{-1}$, which checked well with the parallel-flow value.

As an example of specific surface determination by the parallel-flow method for cotton fibers, the case of a tan cotton of $\frac{7}{8}''$ staple length is taken. A mass of 0.602 g of parallel fiber, filling a tube of 0.542-cm² cross-sectional area and of 2.30-cm length, constituted the sample. This cotton originally was available in the roving form, and consequently the parallel-fiber sample was readily prepared on the Roll-Spacing Determinator. It was found that 224.2 seconds

[5] R. R. Sullivan and K. L. Hertel, Text. Research **11**, 30–38 (1940).

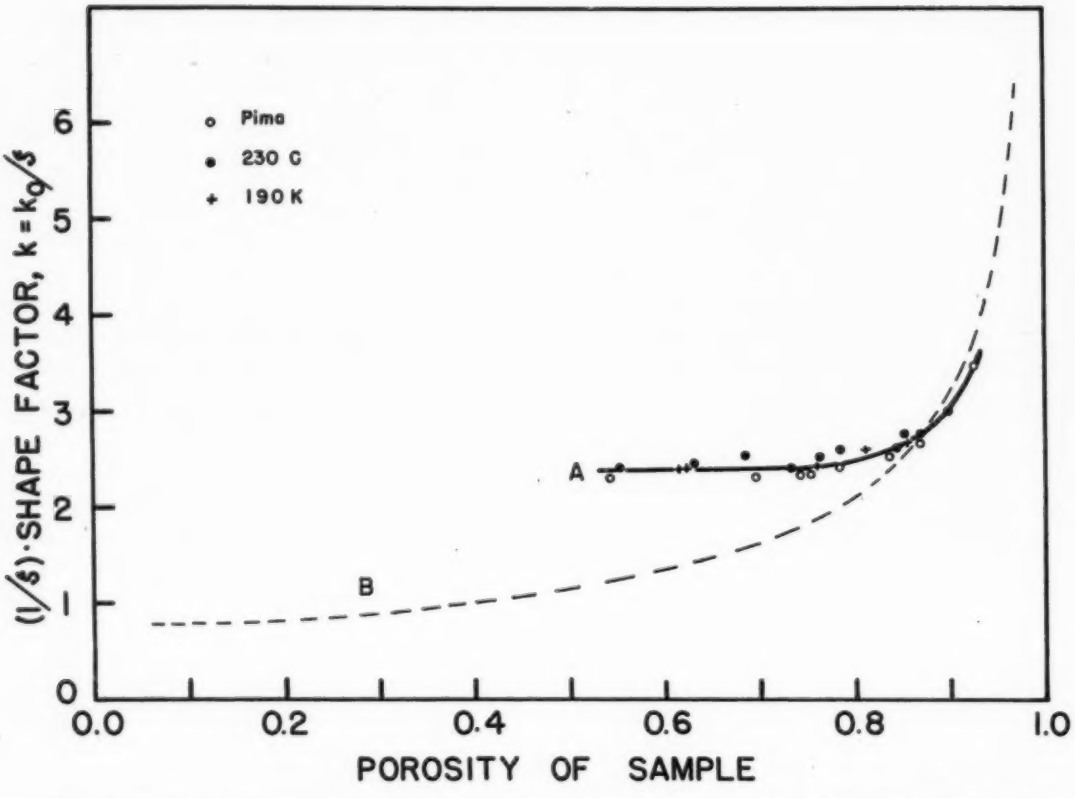



FIG. 2. The relation between the factor $k$ and the porosity $\epsilon$ is given for cases of parallel flow. Curve $A$ is for flow through bundles of parallel cotton fibers. The dashed curve $B$ is for flow through bundles of parallel cylinders.

were required for an air-pressure difference of 1835 dynes/cm² to cause 50 cc of air to flow through the sample. The cotton fiber had a density of 1.54 g/cc. Thus the porosity of the sample was $\epsilon=0.686$. Curve $A$ of Fig. 2 indicates that a value of $k=2.40$ should be used. Recalling that $k=k_0/\zeta$, Eq. (1) may now be used to find $S_0$. Thus a value of $S_0=3.79\times10^3$ cm$^{-1}$ is obtained for the fibers. As a check on this value, the fiber was measured with the arealometer. A value of $S_0=3.68\times10^3$ cm$^{-1}$ was obtained which differed from the parallel-flow method by only 3 percent.

A sample of viscose rayon of porosity $\epsilon=0.548$ was found to have $S_0=2.46\times10^3$ cm$^{-1}$ by the parallel-flow method, using $k_0=1.24$ as obtained from the curve of Fig. 1. A value of $S_0=2.49\times10^3$ cm$^{-1}$ was found for the same fiber by the perpendicular-flow method outlined above for Aralac. There is some evidence, however, to indicate that for rayon which has such highly serrated cross-sectional perimeters the air-flow methods outlined yield values of specific surface which are somewhat lower than those given by a microscope for the same fiber. Further study of air-flow methods of specific surface determinations for rayon fibers should be made, therefore, if accurate absolute values are to be given.


## ACKNOWLEDGMENT

The writer wishes to acknowledge his indebtedness to Professor K. L. Hertel for interest and encouragement, to Miss Louise McCamy and Miss Elizabeth Cotter for aid in the experimental work, to Brookside Mills of Knoxville for use of their Roll-Spacing Determinator, and to the Aralac Corporation of Taftville, Connecticut, for samples of fiber.